

A RAND NOTE

DESCRIPTION OF THE HOUSEHOLD COMPOSITION
AND INCOME FILES CREATED FROM THE INCAP-
RAND GUATEMALAN SURVEY DATA

Carol M. Clark

April 1981

N-1625-RF/AID

Prepared For

The Rockefeller Foundation and
The Agency for International Development



The research described in this report was supported by a grant to the author from the Rockefeller Foundation, and by the Agency for International Development under Contract No. AID/OTR-G-1822/

The Rand Publications Series: The Report is the principal publication documenting and transmitting Rand's major research findings and final research results. The Rand Note reports other outputs of sponsored research for general distribution. Publications of The Rand Corporation do not necessarily reflect the opinions or policies of the sponsors of Rand research.

PREFACE

Initial data collection for the Guatemalan survey was supported by the Rockefeller Foundation and was carried out by the Institute of Nutrition for Central America and Panama (INCAP) in collaboration with The Rand Corporation. Data management and file construction described herein were funded by the Rockefeller Foundation, Ford Foundation, INCAP, and the University of Michigan Department of Economics.

This Note was written with support from the Rockefeller Foundation through a postdoctoral fellowship and a research grant to the author and with support from The Rand Corporation's Family in Economic Development Center. The Center is funded through Grant No. OTR-G-1822 from the Agency for International Development.

The information in this Note facilitates the use of six files containing income, wealth, and household composition data from the INCAP/Rand Guatemalan Survey. The Note may also be of use to researchers who wish to construct their own income, wealth, or household composition variables using the raw data.

SUMMARY

This Note presents a description of the economic and demographic variables contained in the following six data files: a) KID000, b) PAR000, c) CENNUC, d) CENEXT, e) INCOME10, and f) INCEXT. The files were created using the Statistical Analysis System, commonly referred to as "SAS".

The variables contained in these files were created using raw data from the 425 Census and the R10 questionnaire done as part of the INCAP-Rand Guatemalan Survey conducted during 1974-1975.

The KID000 and PAR000 files contain socioeconomic and demographic information on individuals in the communities. The CENNUC and CENEXT files contain household size and composition data on the nuclear and extended families respectively. The income/wealth data for the nuclear and extended families appear in the INCOME10 and INCEXT files respectively.

The Note contains descriptions of each of the variables created. The descriptions detail input variables, formulas, code values and problem cases. It also contains means of the variables and a list of contents of each of the six SAS files.

ACKNOWLEDGMENTS

I wish to thank Albert Anderson and Mike Coble of the Computing Center Staff at the Population Studies Center in Ann Arbor, Michigan, and Pedro Arenales and Peter Russell, in the *Centro de Calculos* at INCAP, for their technical advice and invaluable assistance with the computer facilities at the respective locations. Special thanks go to Albert Anderson at the Population Studies Center, who spent many long hours helping me design and construct early income data files and to William Owens, a statistician at INCAP, who taught me SAS and helped in the early stages of the work as well.

The careful work of Bernard Pillet of INCAP and John Stein at Rand in preparing the economic data for construction of income and wealth measures greatly facilitated my work. Sandra Rosenhouse and Bernard Pillet patiently explained the data management needs and introduced me to resources at INCAP. Marta Amanda Barrera provided excellent information on the communities and original survey procedures, and Miriam Arias lent generous computer assistance and help in reviewing questionnaires. The efforts of Christine Peterson and Mansoor Rashid were valuable in the final cleaning stages.

Terry Fain carefully reviewed the final drafts and offered helpful suggestions.

Kathleen Barden, Sally Dodge, Yasmin Facey, and Virginia Urquhart patiently consulted on layout and expertly typed drafts and the final copy.

Mari Simonen and Georges Jaquemart, and Judith and Frederick Balderston, kindly opened their homes to me while I was in Berkeley. I am also grateful to those members of the Berkeley Project on Education and Nutrition who provided me office space and assistance while I was in Berkeley. This greatly facilitated my work.

Robert Klein of INCAP and William Butz and Dennis DeTray of Rand lent constant support and encouragement of this endeavor.

The data preparation and file construction were funded by the Rockefeller Foundation, the Ford Foundation, INCAP, and the University of Michigan Department of Economics. Preparation of the documentation was supported by the Rockefeller Foundation, The Rand Corporation, and the Agency for International Development. I wish to thank all those affiliated with these funding agencies for their kind assistance in facilitating my work.

Without the generous assistance and support of all of the people mentioned above and many others whom I could not mention here, this project could not have been completed,

CONTENTS

PREFACE iii

SUMMARY v

ACKNOWLEDGMENTS vii

Section

I. INTRODUCTION 1

II. DEMOGRAPHIC AND SOCIOECONOMIC VARIABLES AT THE
INDIVIDUAL LEVEL 7

 2.1 General Comments on the KID000 and PAR000 Files 7

 2.2 Description of Variables Contained in
 KID000 File 8

 2.3 Description of Variables Contained in
 PAR000 File 18

III. DEMOGRAPHIC AND SOCIOECONOMIC VARIABLES AT THE NUCLEAR
FAMILY LEVEL 22

 3.1 General Comments on the CENNUC File 22

 3.2 Description of Variables Contained in
 CENNUC File 23

IV. DEMOGRAPHIC VARIABLES AT THE EXTENDED FAMILY LEVEL 51

 4.1 General Comments on the CENEXT File 51

 4.2 Description of Variables Contained in
 CENEXT File 52

V. INCOME WEALTH AND SOCIOECONOMIC VARIABLES AT THE
NUCLEAR FAMILY LEVEL 59

 5.1 General Comments on the INCOME10 File 59

 5.2 Description of Variables Contained in
 INCOME10 File 61

VI. INCOME AND WEALTH VARIABLES AT THE EXTENDED FAMILY LEVEL 90

 6.1 General Comments on the INCEXT File 90

 6.2 Description of Variables Contained in
 INCEXT File 91

Appendix

A. Notes for CENNUC File 103

B. Notes for CENEXT File 105

C. Notes for INCOME10 and INCEXT Files 106

D. Descriptive Statistics for the Variables Contained
 in the Data Files 107

E. SAS Contents of the Data Files 147

BIBLIOGRAPHY 159

1. . INTRODUCTION

In order to summarize the economic and demographic data obtained in the INCAP-Rand Guatemala Survey, the six data files listed in Table 1 have been created. The purpose of this document is to describe these files in more detail so that they are usable by other researchers interested in using the Guatemala data set.

To facilitate the researcher's understanding and use of the files described in this Note, the INCAP-Rand Guatemala Survey is described briefly and several general comments about the files will be made in this introductory section. Each file and the variables contained in it are described in more detail in the separate sections that follow.

The INCAP-Rand Guatemala Survey

The INCAP-Rand Guatemalan survey is a socioeconomic survey designed by researchers at the Rand Corporation in collaboration with researchers at the Institute for Nutrition in Central America and Panama (INCAP). The fieldwork was done by INCAP personnel between 1974 and 1976.

The sample for the survey consists of all households in four rural and two semi-urban Guatemalan communities. The six communities were chosen to be typical of rural and semi-urban Ladino communities in Central and South America. Virtually all the residents of the six communities are Spanish-speaking.

The relatively isolated rural communities are located between one and two hours driving time (36-102 kilometers) northeast of Guatemala City. Most of the 650-1100 inhabitants are subsistence farmers growing corn and beans. The rural villages are the subject of INCAP's longitudinal study, conducted between 1969 and 1977. This study, funded by NICHD, was designed to investigate the effects of mild to moderate malnutrition on infants' and children's physical and mental development, so longitudinal data exist on children's diet, morbidity, anthropometry and mental development for many of the children in these villages.

Table 1

SUMMARY OF MAJOR ASPECTS OF THE FILES
DESCRIBED IN THIS NOTE

File Name	Focus of Variables	Unit of Observation	Source Questionnaire	Number of Variables	Number of Observations	Type of Data Set	Creation Date	Date of Last Cleaning
KID000	Socioeconomic characteristics of individuals	Individuals with ID number 03-98	425 Census	22	9083	SAS	11/12/79	7/9/80
PAR000	Socioeconomic characteristics of individuals	Individuals with ID number 01-02	425 Census	17	3182	SAS	10/12/79	10/12/79
CENNUC	Household composition and socioeconomic characteristics	Nuclear families	425 Census	78	1832	SAS	11/15/79	7/9/80
CENEXT	Household composition	Extended families	425 Census	18	1558	SAS	11/15/79	11/15/79
INCOME10	Income and wealth and socioeconomic characteristics	Nuclear family	R10 Questionnaire	81	1436	SAS	11/13/79	3/20/80
INCEXT	Income and wealth and socioeconomic characteristics	Extended family	R10 Questionnaire	27	669	SAS	3/26/80	4/30/80

The two semi-urban communities are located just south of Guatemala City, and can be reached in a 20-minute bus ride. Their combined population is about 5,000. The inhabitants are more modern--many are involved in non-agricultural types of income-earning activity and/or commute to the city to work. These villages were not part of the longitudinal study, but some anthropometric and morbidity data on children may become available.

The data for the INCAP-Rand Guatemala survey were collected using 11 survey instruments administered once or several times throughout the survey. Most questionnaires were administered to a large subset of the 1830 households in the sampled villages. Data include: a) men's and women's retrospective life histories, b) home stimulation, c) modernism, d) mother's vocabulary, e) schooling experience, f) women's and children's time use, g) income and wealth, h) men's and women's attitudes and expectations and i) a household census.

File Construction

As indicated in Table 1, the variables contained in the files were created from data obtained in the following two questionnaires: a) The 425 Family Census, and b) R10 Questionnaire on Family Income and Wealth. Hereafter these are called simply the 425 Census and the R10 Survey or questionnaire, respectively.

Unit of Observation: It can be seen from Table 1 that the files exist for different units of observation: the individual, the nuclear family and the extended family. The individual level files provide socioeconomic and demographic data on each individual; they also can be used for creating household composition and other demographic variables describing the nuclear and extended families in the survey.

Several of the files are at the extended family level, a unit of analysis not previously seen in this data set. While much of the data was collected at the nuclear family level, the economic arrangements and informal support systems that exist between related nuclear families in these communities suggested the need to identify "extended" families--families who shared financial resources. Therefore, efforts were made to identify extended families based on interviewer comments and other clues that could be found on the original questionnaires. The CENEXT and INCEXT files were created to summarize the demographic and economic characteristics of these extended families. The CENNUC file provides sufficient data for researchers to determine which nuclear families are in such a resource-sharing arrangement.

The task of determining which families were sharing financial resources was difficult, since no specific questions were asked about it. These files, however, represent the only effort to do this and provide a first step toward specification of extended family units for this data set.

File Creation and Checking: Table 1 indicates the date the current files were created. The preparation for construction of the files took place over a two-and-a-half-year period. Simple data manipulation and cleaning procedures were performed at INCAP in Guatemala. The computer facilities of the University of Michigan and the University of California at San Francisco were used for the more

complicated procedures. Selected consistency checks and searches for out-of-range codes were done during the various stages of file preparation. In addition, attempts to merge parts of these files with other files from the INCAP-Rand survey and with selected longitudinal files and the actual use of these files for analyses turned up some additional inconsistencies in the data. Where questions have arisen, the original questionnaires were consulted.

Where possible, all known problems have been rectified on the files described here. Two problems which could not be corrected due to insufficient information, and two coding errors discovered during the final cleaning of the CENEXT file, remain on the current versions of these files. These problems are described in Figure 1. Other coding errors may turn up as these files are used for analyses. Whenever possible these problems should be verified with the original questionnaires.

Finally, as indicated in Table 1, the files described in this document were created using the Statistical Analysis System, commonly referred to as "SAS". SAS is a computer system for data management and analysis developed at the SAS Institute in Raleigh, North Carolina. The system is particularly good for complex data management tasks and is fairly easy to learn so that a researcher can do much of his/her own computer work. For further information see the SAS User's Guide noted in the References.

Format of the File Descriptions: In the remainder of this Note the particular files and variables contained in them will be described in more detail. Each variable is listed in alphabetical order within the section for the file in which it is contained. For each variable a short six-letter variable name is given in capital letters. This is followed by a longer name which describes the variable more completely. The six-letter name can be used to access the variable on the appropriate SAS data file. Each variable description contains a section detailing: a) input variables, b) coding procedures and, in most cases, c) notes and comments. Since the majority of these variables were constructed from raw data or intermediate data sets,

PAR000.SAS2:

On this version of the PAR000 file there is a problem with PERSON = 01 in NUCFAM 535 in VILAGE = 32. His 425 Census record shows he was born in 1975. This is incorrect, but could not be verified with the original questionnaire. The decision of what to do with this case is left up to the researcher.

CENNUC.SAS5:

On this version of the CENNUC file there is a problem with NUCFAM = 189 in VILAGE = 14. This case has NEMBRZ = 3, NHIJOS = 2, NMUERT = 1 and NMORTI = 1. NHIJOS + NMUERT + NMORTI should equal NEMBRZ. There is no way to tell from the Census 425 files which is correct.

CENEXT.SAS2:

On this version of the CENEXT file the following corrections must be made:

<u>VILAGE</u>	<u>EXTFAM</u>	<u>CORRECTION</u>
06	261	DR006X = 100
08	068	CNT1PX = 1

Figure 1--Miscellaneous Corrections Needed

the variables they were based on are specified under the input variable section. The coding section details code values and construction. Appendixes A, B and C contain several file cleaning notes. Appendix D contains copies of the file directories produced by SAS ("contents"), and Appendix E contains the summary descriptive statistics for each file.

II. DEMOGRAPHIC AND SOCIOECONOMIC VARIABLES AT THE INDIVIDUAL LEVEL

2.1 GENERAL COMMENTS ON THE KID000 AND PAR000 FILES

The KID000 and PAR000 files contain demographic and general socioeconomic variables for *individuals* in the study communities. Each observation corresponds to an individual appearing in cards 101 to 125 of the 425 Census file; all variables in these files were created from these cards. The PAR000 file contains data on individuals with subject identification numbers 01 to 02 (i.e., head and wife of head). The KID000 file contains data on individuals with subject identification numbers 03 to 98.

Due to the way the 425 Census was taken a person may appear *more than once* in the files with different identification numbers (PERSON) and with different codes for status in the household (ESTADO). For example, a son may have left his parents' family to become head of his own family. His record in the KID000 file would have one set of nuclear family and person identification numbers as a child in the family of origin and a code of 4 for ESTADO. He would also appear in the PAR000 file with the family number of the new nuclear family he has formed and a PERSON id of 01 and an ESTADO of 1.

These two files together can be used to construct various household composition variables, including counts of individuals in the nuclear family (e.g. the number of 7-14-year-old males, etc.), dependency ratios, etc. Some such variables are included in the CENNUC file. Researchers may desire others as well. When constructing such variables, care must be taken not to double-count the individuals who appear more than once in the files.

2.2 DESCRIPTION OF VARIABLES CONTAINED IN KID000 FILE

AGEMOS Age in months of individual on March 1, 1975.

Input Variables:

425 Census, 100 series cards, columns 52-57.

Coding:

The value is the age in months of the individual on March 1, 1975.

Notes and Comments:

If the month of birth was missing, it was assumed to be March. If the year was missing, no assumption was made and AGEMOS was set to missing. The exception to this is individual 31 068 84 whose AGEMOS was assumed to be 720, since this was the mother of the female head of household.

AGOLDR Age in months of child's next older sibling listed in 425 Census.

Input Variables:

425 Census, 100 series cards, columns 52-57.

Coding:

The value is the age in months of the individual's next older sibling who was listed in the 425 Census.

Notes and Comments:

AGOLDR is coded as missing if: a) child's older sibling has missing data on month and/or year of birth, b) child in question is the oldest child listed in the nuclear family, or c) individual is not a biological child of the person listed as 02 in the nuclear family.

AGYNGR Age in months of child's next younger sibling listed in 425 Census.

Input Variables:

425 Census, 100 series cards, columns 52-57.

Coding:

The value is the age in months of the individual's next younger sibling who was listed in the 425 Census.

Notes and Comments:

AGYNGR is coded as missing if: a) child's younger sibling has missing data on month and/or year of birth, b) child in question is the youngest child listed in the nuclear family, or c) individual is not a biological child of the person listed as 02 in the nuclear family.

ALFAKI Index of individual's literacy.

Input Variables:

425 Census, 100 series cards, columns 73.

Coding:

0 does not read and write
1 reads and writes with difficulty
2 reads and writes without difficulty

Notes and Comments:

The information was obtained by asking the respondent. No test was made to see if the individual in question actually could read when 1 or 2 was coded. In the urban areas there may have been some tendency to claim to be literate even when the respondent was not, since jobs, etc., are more likely to be contingent on one's ability to read and write.

ANONAC Birth year of individual.

Input Variables:

425 Census, 100 series cards, columns 56-57.

Coding:

ANONAC is the last two digits of the birth year of the individual.

Notes and Comments:

In order to obtain complete birth date, ANONAC must be used in conjunction with DAYNAC AND MESNAC.

BIOLDR Birth interval to next older sibling.

Input Variables:

425 Census, 100 series cards, columns 52-57.

Coding:

The value is the actual number of months between the birth of the child and the birth of the next older sibling listed in the 425 Census.

Notes and Comments:

BIOLDR is coded as missing if: a) child's older sibling has missing data on month and/or year of birth, (See Note 2), b) child has missing data on month and/or year of birth, c) child is the oldest child listed or d) individual is not a biological child of the person listed as 02 in the nuclear family.

BIRTHO

Birth Order.

Input Variables:

425 Census, 100 series cards, columns 52-57.

Coding:

The value is the birth order of the child including all children who were listed in the 425 Census and who were born alive or dead (ESTADO =1-6) to person with person number 02 in the nuclear family.

BIYNGR

Birth interval to next younger sibling.

Input Variables:

425 Census, 100 series cards, columns 52-57.

Coding:

The value is the actual number of months between the birth of the child and the birth of the next younger sibling listed in the 425 Census.

Notes and Comments:

BIYNGR is coded as missing if a) younger sibling has missing data on month and/or year of birth b) individual has missing data on month and/or year of birth, c) individual is the youngest child listed, or d) individual is not a biological child of 02.

DAYNAC

Day of birth of individual.

Input Variables:

425 Census, 100 series cards, columns 52-53.

Coding:

DAYNAC is the number of the day; ranging from 01 to 31.

Notes and Comments:

In order to obtain complete date of birth DAYNAC must be used in conjunction with MESNAC and ANONAC.

ESOLDR Current living status of next older sibling who is listed in the 425 Census.

Input Variables:

425 Census, 100 series cards, column 46.

Coding:

See ESTADO below.

Notes and Comments:

ESOLDR is coded as missing if a) individual is the oldest child in the nuclear family, or b) individual is *not* a biological child of 02. ESOLDR is a 2 digit code if next older siblings were twins. In this case, first digit reflects current status of one of twins and second digit reflects current status of the other.

ESTADO Current living status of individual.

Input Variables:

425 Census, 100 series cards, column 46.

Coding:

- 1 living in and forming part of nuclear family
- 2 dead
- 3 living in house of nuclear family, but are considered as part of another nuclear family
- 4 living out of house of nuclear family as part of another nuclear family
- 5 living outside of the community
- 6 stillborn

Notes and Comments:

Due to the data collection procedure, a person may appear more than once in the file with different ID numbers and with a code

of ESTADO = 1 in one place and a code of ESTADO = 3 or 4 in the other.

ESYNGR Current living status of next younger sibling who is listed in the 425 Census.

Input Variables:

425 Census, 100 series cards, column 46.

Coding:

See ESTADO below.

Notes and Comments:

ESYNGR is coded as missing if a) the child is the youngest child in the nuclear family or b) individual is *not* a biological child of 02. ESYNGR is a 2-digit code if next youngest siblings were twins. In this case first digit reflects current status of one of twins and second digit reflects current status of other.

GRADOK Number of school grades completed by individual.

Input Variables:

425 Census, 100 series cards, columns 74-75.

Coding:

GRADOK is the actual number of school grades completed by individual.

LBRTHO Live birth order.

Input Variables:

425 Census, 100 series cards, columns 52-57.

Coding:

The value is the live birth order of the child including all children who were listed in the 425 Census and who were born alive (ESTADO = 1-5) to person with person number 02 in the nuclear family.

MESNAC Month of birth of individual.

Input Variables:

425 Census, 100 series cards, columns 54-55.

Coding:

MESNAC is the number of the month in which the individual was born, ranging from 01 to 12.

Notes and Comments:

In order to obtain complete birth date MESNAC must be used in conjunction with DAYNAC and ANONAC.

NAGEMS

Number of individuals in nuclear family with missing data on age.

Input Variables:

425 Census, 100 series cards, columns 52-57.

Coding:

NAGEMS is the actual number of individuals listed in nuclear family's 425 Census who had missing birth date data.

NUCFAM

Nuclear family identification number.

Input Variables:

425 Census, 100 series cards, columns 9-11.

Coding:

NUCFAM is the nuclear family ID number assigned to the family at the beginning of the study.

Notes and Comments:

In order to obtain an identification number which is distinct from the ID numbers of all other individuals, this variable must be used in conjunction with VILAGE and PERSON.

OCACTK

Current occupation of individual.

Input Variables:

425 Census, 100 series cards, columns 79-80.

Coding:

- 00 Domestic duties around own house, less than 10 years old, not applicable.
- 01 Agricultural day worker employed by individuals or agricultural enterprise within the community (works by day or paid by task)

- 02 Agricultural day worker who is employed by individual or agricultural enterprise outside the community
- 03 Agricultural worker who rents land (s)he farms (rent may be paid in cash, in kind or in work)
- 04 Agricultural worker who works land owned by relatives (extended family)
- 05 Agricultural worker who works own land
- 06 Non-agricultural day laborer working within the community, employed by an individual, a private enterprise or government agency dedicated to the fabrication of products. These are manual laborers as distinguished from those who handle machinery (e.g., factory worker, guardian, porter, road worker, truckman's helper, janitor).
- 07 Non-agricultural day laborer working *outside* the community. See 06 above for further detail
- 08 Domestic worker engaged in domestic housekeeping tasks for remuneration. Work can be done either in own home (e.g., taking in laundry, making tortillas for sale to neighbors or local construction crew) or in the home or enterprise of others (e.g., servant, nurse maid, gardener, waiter, etc.)
- 09 Worker in family enterprise producing "traditional" goods or crafts (e.g., weaver, potter, rope maker, hat maker, candlemaker, soap maker, etc.)
- 10 Vendor of family's agricultural production. Must leave home to qualify for this category.
- 11 Seller of crafts or goods produced in family enterprise (e.g., sells tortillas in Guatemala City). Must leave home to qualify for this category.
- 12 Permanent merchant within the community (e.g., store, canteen, pharmacy, butcher shop)
- 13 Merchant who buys goods inside or outside the community and sells them outside the community (e.g., peddler of clothes, or of oranges and melons). Also includes truck drivers.
- 14 Skilled self-employed worker (e.g., tailor, cobbler, builder, seamstress, barber, baker, midwife)
- 15 Employee in a factory or workshop or on construction (one who is contracted to work (e.g., gasoline station attendant, shoe maker's helper, etc.)
- 16 Non-office specialized employee (e.g., policeman, soldier, dispatcher, etc.)
- 17 Office or white collar worker (e.g., teacher, mayor,

secretary, etc.)

- 18 Professional who holds a university degree and who is self employed or affiliated with an institution (e.g., medical doctor, dentist, lawyer, engineer, etc.)
- 19 Owner of large farm or plantation devoted to the production of commercial crops (such as bananas, coffee, or cotton) who hires others to work the land
- 20 Unemployed but looking for work. Individual in this category must be older than 10 years of age unless the child is not in school and does not help the parents with chores. Also includes workers who have no work in the seasons when they normally have work. Individuals whose occupations involve seasonal work are not considered unemployed during the slack seasons.
- 21 Chronically ill, (e.g., tuberculosis, mental retardation or physically disabled)
- 22 Formally retired from work. May or may not receive pension
- 23 Laid off or quit working with severance pay
- 24 Due to old age does not realize usual occupation
- 25 In prison
- 26 Student--an individual older than ten years of age who studies during the day
- 27 Farms communal land

OCPRNK Principal (usual) occupation of individual.

Input Variables:

425 Census, 100 series cards, columns 77-78.

Coding:

See OCACTK above.

PERSON Individual's identification number.

Input Variables:

425 Census, 100 series cards, columns 12-13.

Coding:

03-20 Children of 01 *and* 02

21-30 Children of 01 who form part of the nuclear family

31-40 Children of 02 and *not* of 01 who form part of the nuclear family (all who are living presumably were

- listed)
- 41-50 Nieces and nephews of 01 who form part of the nuclear family
 - 51-60 Nieces and nephews of 02 who form part of the nuclear family
 - 61-70 Grand children who live with and depend on the nuclear family
 - 71-80 Adopted children
 - 81 Father of 01
 - 82 Mother of 01
 - 83 Father of 02
 - 84 Mother of 02
 - 85-90 Other relatives who form part of the nuclear family
 - 91-98 Other non-related people (generally temporarily in the household)

Notes and Comments:

In order to obtain an ID number which is distinct from ID's of all other individuals this variable must be used in conjunction with VILAGE and NUCFAM.

SEXCEN Sex of individual as recorded on 425 Census.

Input Variables:

425 Census, 100 series cards, column 14.

Coding:

1 male
2 female

TWIN10 Index of status of twin.

Input Variables:

425 Census, 100 series cards, columns 52-57.

Coding:

0 if child is not a twin
1 if child is a twin
. if individual is not a biological child of 02.

VILAGE Identification number of village in which individual lives.

Input Variables:

425 Census, 100 series cards, columns 7-8.

Coding:

03 Santo Domingo

06 Conacaste

08 El Espíritu Santo

14 San Juan

31 Colonias San Antonio and El Cortijo

32 San Miguel Petapa

2.3 DESCRIPTION OF VARIABLES CONTAINED IN PAR000 FILE

AGEMOS Age in months of individual on March 1, 1975.

Input Variables:

425 Census, 100 series cards, columns 52-57.

Coding:

See AGEMOS in KID000 file.

Notes and Comments:

See AGEMOS in KID000 file. In addition it should be noted that AGEMOS for individual 32 535 01 was assumed to be 551, since this was the head of the household.

ALFAMA Index of literacy of 02.

Input Variables:

425 Census, 100 series cards, column 73.

Coding:

See ALFAKI in KID000 file for codes.

ALFAPA Index of literacy of 01.

Input Variables:

425 Census, 100 series cards, column 73.

Coding:

See ALFAKI in KID000 file for codes.

ANONAC Birth year of individual.

Input Variables:

425 Census, 100 series cards, columns 56-57.

Coding:

See ANONAC in KID000 file.

DAYNAC Day of birth of individual.

Input Variables:

425 Census, 100 series cards, columns 52-53.

Coding:

See DAYNAC in KID000 file.

ESTADO Current living status of individual

Input Variables:

425 Census, 100 series cards, column 46.

Coding:

- 1 Present--all those who are living in the house at the time of the interview or who have migrated only temporarily to do agricultural work elsewhere (e.g., to the coast to harvest cotton). Also includes husbands or wives who are away from the community but maintain direct links with family by visiting regularly, and continuing to have direct responsibility for family (e.g., works in Guatemala City during week as policeman and returns on weekends to village). This does *not* include those who live in another country or who have established another household elsewhere.
- 2 Dead
- 3 Live in house but are considered part of another nuclear family
- 4 Live away from the house and form part of another nuclear family
- 5 Live away from community--have made independent life for selves
- 6 Stillborn--those born dead between 28 and 38 weeks of pregnancy
- 7 Emigrated definitively (in some places noted).

GRADOM Number of school grades completed by 02.

Input Variables:

425 Census, 100 series cards, columns 74-75.

Coding:

See GRADOK on KI0000 file.

GRADOP Number of school grades completed by 01.
Input Variables:
425 Census, 100 series cards, columns 74-75.
Coding:
See GRADOK on KID000 file.

MESNAC Month of birth of individual.
Input Variables:
425 Census, 100 series cards, columns 54-55.
Coding:
See MESNAC in KID000 file.

NUCFAM Nuclear family identification number.
Input Variables:
425 Census, 100 series cards, columns 54-55.
Coding:
See NUCFAM in KID000 file.

OCACTM Current occupation of 02.
Input Variables:
425 Census, 100 series cards, columns 79-80.
Coding:
See OCACTK in KID000 file.

OCACTP Current occupation of 01.
Input Variables:
425 Census, 100 series cards, columns 79-80.
Coding:
See OCACTK in KID000 file.

OCPRNM Principal occupation of 02.

Input Variables:

425 Census, 100 series cards, columns 77-78.

Coding:

See OCACTK in KID000 file.

OCPRNP Principal occupation of 01.

Input Variables:

425 Census, 100 series, columns 77-78.

Coding:

See OCACTK in KID000 file.

PERSON See PERSON in KID000 file for codes.

Input Variables:

425 Census, 100 series cards, columns 12-13.

Coding:

01 male head
02 female head, (or wife of 01)

Notes and Comments:

In order to obtain an ID number which is distinct from ID numbers of all other individuals, this variable must be used in conjunction with VILAGE and NUCFAM.

SEXCEN Sex of individual as recorded on 425 Census.

Input Variables:

425 Census, 100 series cards, column 14.

Coding:

See SEXCEN in KID000 file.

VILAGE Identification number of village in which individual lives

Input Variable:

425 Census, 100 series cards, columns 7-8.

Coding:

See VILAGE in KID000 file.

III. DEMOGRAPHIC AND SOCIOECONOMIC VARIABLES AT THE NUCLEAR FAMILY LEVEL

3.1 GENERAL COMMENTS ON THE CENNUC FILE

The CENNUC file contains demographic and socioeconomic data on the *nuclear families* in the study communities. Each observation in the CENNUC file corresponds to a nuclear family who was surveyed in the 425 Census. Some of the variables appearing in this file were taken from cards 000 and 001 of the 425 Census. The household size and composition variables were constructed based on the KID000 and PAR000 files.

The CENNUC file also identifies nuclear families who live together and/or share income (here called extended families). Because it contains the extended family identification number for each nuclear family in the survey (EXTFAM), the CENNUC file serves as an interface between the individual level files (KID000 and PAR000) and the extended family files CENEXT and INCEXT.

3.2 DESCRIPTION OF VARIABLES CONTAINED IN CENNUC FILE

ADDRES Address of nuclear family.

Input Variables:

425 Census, card 000, columns 22-26 .

Coding:

ADDRES is formed from the block number, zone number and house number in which the nuclear family resides at the time of the 425 Census.

AGEM01 Age of 01 in months.

Input Variables:

425 Census, 100 series cards, columns 52-57.

Coding:

AGEM01 is the actual number of months from birth until March 1975.

AGEM02 Age of 02 in months.

Input Variables:

425 Census, 100 series cards, columns 52-57.

Coding:

Actual number of months from birth until March 1975.

AGEMIS Index of missing birth date data for any family member who was reported as currently living in nuclear family (e.g., ESTADO=1) at time of the 425 Census.

Input Variables:

425 Census, 100 series cards, columns 52-57.

Coding:

- ° nuclear family was given an address but members were included in the census of another family. See Note 1 in Appendix A.
- 0 no missing birth date data on any member listed as currently living with the nuclear family.
- 1 missing birth date data on one or more members listed as currently living with the nuclear family.

Notes and Comments:

This variable was constructed for data management purposes and to enable researchers to know what underlies the following count variables: FB0000--FB60UP, FT0000--FT60UP, MB0000--MB60UP, MT0000--MT60UP. Since occasionally birth dates were not known, those nuclear families for whom AGEMIS = 1 will exclude one or more people from the counts. Generally, missing data appears for grandparents and/or more distant relatives (e.g., cousins, nieces, uncles, etc.) who live with the nuclear family.

AGEMSF Index of missing birth date data for any nuclear family member.

Input Variables:

PAR000 file: ESTADO, AGEMOS
KID000 file: ESTADO, AGEMOS

Coding:

AGEMSF is a tally of people in the nuclear family at the 425 Census with missing data on birth date. A missing data code was assigned when a nuclear family had been given an address but members were included in the 425 Census of another nuclear family. See Note 1 in Appendix A.

Notes and Comments:

This variable was constructed to enable the researcher to know what underlies the following nuclear family count variables: DENOMR, NUM006, NUM714, EACONF, DRO06F DR714F. Since occasionally birth dates were not known, those nuclear families for whom AGEMSF > 0 will exclude one or more people from the count. It is rare that young biological children of 02 have missing birth date data. Generally missing data appears for grandparents and/or more distant relatives (e.g., cousins, nieces, uncles, etc.) who live with the nuclear family.

ALFA01 Index of literacy of person 01.

Input Variables:

PAR000 file: ALFAPA

Coding:

0 Does not know how to read and write
1 Reads and writes with difficulty
2 Reads and writes without difficulty

Notes and Comments:

See ALFAKI in the KID000 file for additional comments on collection of literacy data

ALFA02 Index of literacy of person 02

Input Variables:

PAR000 file: ALFAMA

Coding:

- 0 Does not know how to read and write
- 1 Reads and writes with difficulty
- 2 Reads and writes without difficulty

Notes and Comments:

See ALFAKI in the KID000 file for additional comments.

CNTPIF Number of people currently living in the nuclear family.

Input Variables:

KID000 file: AGEMOS, ESTADO

PAR000 file: AGEMOS, ESTADO

Coding:

CNTPIF is the sum of people listed as currently living in the nuclear family (e.g., ESTADO = 1) at the time of the 425 Census.

Notes and Comments:

All people with ID's 01-90 were counted regardless of whether or not they had missing age data. A code of "0" appears for those families listed in Note 1, Appendix A. See also Note 2, Appendix A for treatment of families who migrated in 1974-1975.

The value of CNTPIF does not always equal the value of NPERFM since respondents' reports often include those who had an ESTADO other than 1 or those with a person ID number greater than 90.

CNTPRF Number of people reported as currently living in nuclear family for whom birth date data were available.

Input Variables:

KID000 file: AGEMOS, ESTADO
PAR000 file: AGEMOS, ESTADO

Coding:

CNTPRF is the sum of people reported as currently living in the nuclear family (e.g., ESTADO = 1) at the time of the 425 Census.

Notes and Comments:

All people with IDs 01-90 who were reported as present in nuclear families were counted. However, unlike CNTPIF, only those for whom age in months could be calculated were included in the count. A code of "0" appears for those families listed in Note 1, Appendix A. See also Note 2, Appendix A for families who migrated in 1974-1975.

DADPRS Index of presence of person 01 in nuclear family.

Input Variables:

PAR000 file: PERSON, ESTADO

Coding:

- 1 If ESTADO = 1 for person number 01
- 1 otherwise (i.e. ESTADO = 2 - 5)
 - 425 Census card for 01 missing from file. See Note 3, Appendix A.

DAYCEN Day the 425 Census was taken.

Input Variables:

425 Census, card 000, columns 18-19

Coding:

DAYCEN is the number of the day in the month that the 425 Census was taken ranging from 01 to 31.

Notes and Comments:

In order to obtain complete date of census DAYCEN must be used in conjunction with MONTHC and YEARCE.

DENOMR Number of adults aged 15-60 currently living in the nuclear family.

Input Variables:

KID000 file: AGEMOS, ESTADO
PAR000 file: AGEMOS, ESTADO

Coding:

DENOMR is the sum of adults in the nuclear family for whom
PERSON = 01-90, 180 ≤ AGEMOS ≤ 719 and ESTADO = 1.

Notes and Comments:

Count includes only those for whom age in months could be
calculated. See AGEMSF to determine which nuclear families
have members with missing birth date data.

DR006F Number of 0-6 year olds per adults (aged 15-60) in nuclear
family.

Input Variables:

CENNUC file: NUM006, DENOMR

Coding:

DR006F is the value of the following:

$$(NUM006/DENOMR) * 100$$

Notes and Comments:

Components NUM006 and DENOMR include only those individuals
for whom age in months could be calculated. See AGEMSF to
determine which nuclear families have numbers with missing
birth date data.

DR714F Number of 7-14 year olds per adult (aged 15-60) in nuclear
family.

Input Variables:

CENNUC file: NUM714, DENOMR

Coding:

DR714F is the value of the following:

$$(NUM714/DENOMR) * 100$$

Notes and Comments:

Components NUM714 and DENOMR include only those individuals
for whom age in months could be calculated. See AGEMSF to
determine which nuclear families have members with missing
birth date data.

EACONF Number of equivalent adult consumers in nuclear family.

Input Variables:

KID000 file: AGEMOS, ESTADO
PAR000 file: AGEMOS, ESTADO

Coding:

EACONF is a weighted sum of the number of people with ID numbers 01-90 present in the household (ESTADO = 1) where the weights are based on nutrient requirements of different ages and sexes relative to an adult male. EACONF is the value of the following:

$$\left(\sum_{i=1}^n (\text{AGEMOS}_i) \text{WGT}_i \right) * 100$$

where AGEMOS_i = age in months of individual i in the household ($i = 1, 2, \dots, n$)

WGT_i = weight assigned to individuals of i 's age and sex

n = number of individuals in the household

Weights used are shown in Table 2 below.

Table 2

WEIGHTS USED IN CALCULATING EACONF

	Male	Female
0-8 months	.33	.33
9-11 months	.36	.36
1 year	.40	.40
2 years	.47	.47
3 years	.53	.53
4-6 years	.60	.60
7-9 years	.71	.71
10-12 years	.86	.78
13-15 years	.98	.84
16-19 years	1.07	.79
19-40 years	1.00	.71
40+ years	.89	.67

Notes and Comments:

These weights were based on figures published by INCAP (INCAP, 1973). Since these apply to Central America, it was thought that they would most adequately take into account the body size, activity level, climate, and other factors peculiar to Guatemala which affect calorie requirements. As seen from Table 3, INCAP's figures are similar to those estimated by the FAO (Sukhatme, 1976).

Table 3

CALORIC REQUIREMENTS BY AGE AND SEX RELATIVE TO MALE ADULT AGED 19-40
AS CALCULATED BY INCAP, FAO, AND NATIONAL ACADEMY OF SCIENCES

	INCAP		FAO		RDA	
	Male	Female	Male	Female	Male	Female
0-8 months	.3345	.3345	.2733	.2733		
9-11 months	.3552	.3552				
1 year	.3966	.3966	.4533	.4533	.4815	.4815
2 years	.4655	.4655				
3 years	.5345	.5345				
4-6 years	.6034	.6034	.6100	.6100	.6667	.6667
7-9 years	.7069	.7069	.7300	.7300		
7-10 years					.8889	.8889
10-12 years	.8621	.7759	.8667	.7833		
11-14 years					1.037	.8889
13-15 years	.9828	.8448	.9667	.8300		
15-22 years					1.1111	.7778
16-18 years	1.069	.7931				
16-19 years			1.0233	.7700		
19-40 years	1.00	.7069	1.00	.7333		
23-50 years					1.00	.7407
40+ years				.8650		
51+ years					.8889	.6667

Both INCAP's and FAO's publications indicate a decline in calorie requirements with age. There is some ambiguity as to when the decline occurs and the magnitude of the decline. I assumed that activity level, and thus calorie requirements, decline once around 50 and remain at the new level thereafter, since people in the socioeconomic group included in the sample generally tend to work until they die. In the

absence of more concrete numbers for Central America, I based the relative consumption requirements after 50 years of age on the Recommended Daily Dietary Allowance estimated by the Food and Nutrition Board of the National Academy of Sciences (Church, 1975.) for U.S. populations. This was felt to be satisfactory for several reasons: 1) We are interested in relative calorie requirements, not absolute ones. Absolute calories requirements differ somewhat between Latin America and U.S. but as seen in Table 2, requirements relative to adult males in either context differ much less. 2) They are similar to some rough estimates by FAO. 3) They are the most recent figures available.

Unfortunately, no allowance could be made for pregnant and lactating women because it was not possible from the available data to tell the reproductive status of the women.

ESTCIV Civil Status of head of household.

Input Variables:

425 Census, card 000, column 27

Coding:

- 1 Single
- 2 Single without formal union
- 3 Single mother
- 4 Mother without stable union
- 5 Consensual union
- 6 Married
- 7 Separated or divorced
- 8 Widow

ETHNIC Ethnic group of head of household.

Input Variables:

425 Census, card 000, column 28

Coding:

- 1 Indian
- 2 Spanish-speaking

EXTFAM Extended family identification number.

Input Variables:

R10 survey and 425 Census, interviewer observations, and field supervisor notes.

Coding:

EXTFAM is an identification number for the extended family.

Notes and Comments:

In these villages several nuclear families may live together and share income or they may live apart but still share income. Usually such families are related (e.g., brothers, parents and adult sons, etc.). This variable allows the researcher to relate the nuclear families in this arrangement to one another. In order to distinguish any extended family from all others. The variable EXTFAM must be used in conjunction with VILAGE. See HHFORM in this file for more information on extended family relations and coding.

EXTNUM Identification number of nuclear family in extended family.

Input Variables:

CENNUC file: EXTFAM

Coding:

- 1 first nuclear family in extended family encountered when this variable was created
- 2 second such family
- 3 third such family
- 4 fourth such family

Notes and Comments:

The number assigned has no particular sociological demographic, or economic significance. It is merely a variable created to distinguish one nuclear family from another within the extended family.

FB0000 Number of female children of 02 who are less than 1 year old and currently living with the nuclear family.

Input Variables:

KID000 file: ESTADO, AGEMOS, SEXCEN, PERSON

Coding:

FB0000 is the sum of children in the nuclear family with

PERSON = 03-21 or 31-40 for whom SEXCEN = 2, AGEMOS < 12 and ESTADO = 1.

Notes and Comments:

See Note 4 in Appendix A.

FB0106 Number of female children of 02 who are 1-6 years old and currently living with the nuclear family.

Input Variables:

KID000 file: ESTADO, AGEMOS, PERSON, SEXCEN

Coding:

FB0106 is the sum of children in the nuclear family.
PERSON = 03-21 or 31-40 for whom SEXCEN = 2, 12 ≤ AGEMOS < 83 and ESTADO = 1.

Notes and Comments:

See Note 4 in Appendix A.

FB0710 Number of female children of 02 who are 7-10 years old currently living with the nuclear family.

Input Variables:

KID000 file: ESTADO, AGEMOS, PERSON, SEXCEN

Coding:

FB0710 is the sum of children in the nuclear family with
PERSON = 03-21 or 31-40 for whom SEXCEN = 2, 84 ≤ AGEMOS < 131 and ESTADO = 1.

Notes and Comments:

See Note 4 in Appendix A.

FB1114 Number of female children of 02 who are 11-14 years old and currently living with the nuclear family.

Input Variables:

KID000 file: ESTADO, AGEMOS, PERSON, SEXCEN

Coding:

FB1114 is the sum of children in the nuclear family with
PERSON = 03-21 or 31-40 for whom SEXCEN = 2, 132 ≤ AGEMOS < 179 and ESTADO = 1.

Notes and Comments:

See Note 4 in Appendix A.

FB1545 Number of female children of 02 who are 15-45 years old and currently living with the nuclear family.

Input Variables:

KID000 file: ESTADO, AGEMOS, PERSON, SEXCEN

Coding:

FB1545 is the sum of children in the nuclear family with PERSON = 03-21 or 31-40 for whom SEXCEN = 1, $180 \leq$ AGEMOS \leq 551 and ESTADO = 1.

Notes and Comments:

See Note 4 in Appendix A.

FB4659 Number of female children of 02 who are 46-59 years old and currently living with the nuclear family.

Input Variables:

KID000 file: ESTADO, AGEMOS, PERSON, SEXCEN

Coding:

FB4659 is the sum of children in the nuclear family with PERSON = 03-21 or 31-40 for whom SEXCEN = 2, $552 \leq$ AGEMOS \leq 719, and ESTADO = 1.

Notes and Comments:

See Note 4 in Appendix A.

FB60UP Number of children of 02 who are 60 years and older and currently living with the nuclear family.

Input Variables:

KID000 file: ESTADO, AGEMOS, PERSON, SEXCEN

Coding:

FB60UP is the sum of children in the nuclear family with PERSON = 03-21 or 31-40 for whom SEXCEN = 2, AGEMOS \geq 720 and ESTADO = 1.

Notes and Comments:

See Note 4 in Appendix A.

FT0000 Number of related females who are less than 1 year old and currently living with the nuclear family.

Input Variables:

KID000 file: ESTADO, AGEMOS, PERSON, SEXCEN

PAR000 file: ESTADO, AGEMOS, PERSON, SEXCEN

Coding:

FT0000 is the sum of nuclear family members with PERSON = 01-90 for whom SEXCEN = 2, AGEMOS \leq 12 and ESTADO = 1.

Notes and Comments:

See Note 5 in Appendix A.

FT0106 Number of related females who are 1-6 years old and currently living with the nuclear family.

Input Variables:

KID000 file: ESTADO, AGEMOS, PERSON, SEXCEN

PAR000 file: ESTADO, AGEMOS, PERSON, SEXCEN

Coding:

FT0106 is the sum of nuclear family members with PERSON = 01-90 for whom SEXCEN = 2, $12 \leq$ AGEMOS \leq 83, and ESTADO = 1.

Notes and Comments:

See Note 5 in Appendix A.

FT0710 Number of related females who are 7-10 years old and currently living with the nuclear family.

Input Variables:

KID000 file: ESTADO, AGEMOS, PERSON, SEXCEN

PAR000 file: ESTADO, AGEMOS, PERSON, SEXCEN

Coding:

FT0710 is the sum of nuclear family members with PERSON = 01-90 for whom SEXCEN = 2, $84 \leq$ AGEMOS \leq 131 and ESTADO = 1.

Notes and Comments:

See Note 5 in Appendix A.

FT1114 Number of related females who are 11-14 years old and currently living with the nuclear family.

Input Variables:

KID000 file: ESTADO, AGEMOS, PERSON, SEXCEN
PAR000 file: ESTADO, AGEMOS, PERSON, SEXCEN

Coding:

FT1114 is the sum of nuclear family members with PERSON = 01-90 for whom SEXCEN = 2, $132 \leq \text{AGEMOS} \leq 179$ and ESTADO = 1.

Notes and Comments:

See Note 5 in Appendix A.

FT1545 Number of related females who are 15-45 years old and currently living with the nuclear family.

Input Variables:

KID000 file: ESTADO, AGEMOS, PERSON, SEXCEN
PAR000 file: ESTADO, AGEMOS, PERSON, SEXCEN

Coding:

FT1545 is the sum of nuclear family members with PERSON = 01-90 for whom SEXCEN = 2, $180 \leq \text{AGEMOS} \leq 551$ and ESTADO = 1.

Notes and Comments:

See Note 5 in Appendix A.

FT4659 Number of related females who are 46-59 years old and currently living with the nuclear family.

Input Variables:

KID000 file: ESTADO, AGEMOS, PERSON, SEXCEN
PAR000 file: ESTADO, AGEMOS, PERSON, SEXCEN

Coding:

FT4659 is the sum of nuclear family members with PERSON = 01-90 for whom SEXCEN = 2, $552 \leq \text{AGEMOS} \leq 719$ and ESTADO = 1.

Notes and Comments:

See Note 5 in Appendix A.

FT60UP Number of related females who are 60 years old and older and currently living with the nuclear family.

Input Variables:

KID000 file: ESTADO, AGEMOS, PERSON, SEXCEN
PAR000 file: ESTADO, AGEMOS, PERSON, SEXCEN

Coding:

FT6OUP is the sum of nuclear family members with PERSON = 01-90 for whom SEXCEN = 2, AGEMOS \geq 720 and ESTADO = 1.

Notes and Comments:

See Note 5 in Appendix A.

GRAD01 Grades attained by 01.

Input Variables:

PAR000 file: GRADOP

Coding:

GRAD01 is the actual number of school grades completed by 01.

GRAD02 Grades attained by 02.

Input Variables:

PAR000 file: GRADOM

Coding:

GRAD02 is the actual number of school grades completed by 02.

HHFORM Type relation household has to other related families in community.

Input Variables:

R10 survey and 425 Census, interviewer observations, and field supervisor notes.

Coding:

- 0 Nuclear family living alone
- 1 Nuclear family living at same address as another nuclear family but does not appear to constitute an extended family and/or to share income (e.g., roomers, apartment dwellers, etc.)
- 2 Emigrated from a community. (Another family may now be living at address.)
- 3, 5 & 7 Extended family composed of two or more nuclear families living at same address. 5 and 7 indicate possible missing income data

- 4 & 6 Income sharing between two or more nuclear families who do not live at same address. 6 indicates possible missing income data.
- 8 Two or more nuclear families sharing income but we do not know who other family(ies) is(are).

Notes and Comments:

No specific question was asked on any questionnaire about nuclear families sharing resources with other nuclear families [usually such families are related (e.g., brothers, sons, etc.)], nor was there a question specifically identifying which families were in such an arrangement. Therefore, the relation of nuclear families to each other had to be established by reviewing the 425 Census forms and the R10 questionnaires. The work was based initially on a match of addresses on the 425 Census. Field supervisor's record books in which relevant observations were often recorded were also reviewed. In order to be considered in an extended family/income-sharing arrangement a) the relation had to be stated specifically in interviewers' observations, b) the individual had to indicate living with relatives (425 Census, card 001, column 24) and the description of the houses as given on card 001 of the 425 Census had to jibe or c) the individual had to indicate sharing of agricultural work or income from a family farm or enterprise (R10 form). In cases where there was doubt, families were assumed *not* to be in an extended family/income-sharing relation.

For purposes of this variable, a family in which the mother and/or father of 01 or 02 were listed as members of a nuclear family on the 425 Census (e.g., had ID numbers 81-84), they were not considered part of a separate nuclear family. This household received code of 0 on HHFORM. If separate 425 Census existed for the mother and/or father of 01 or 02, but review of the questionnaires indicated that they lived in the same house, then code value of 3 for HHFORM was given. While there is essentially no difference in these two situations, it would have made the coding and data management even more complicated than it already was. Since these efforts to discover income-sharing arrangements between nuclear families were done primarily to obtain a more accurate estimate of economic resources available per person in the unit that was sharing income, it was not deemed sufficiently useful for the current purposes to distinguish the two types of households.

Finally, it should be noted that there may be some error in the coding of whether or not *income data do exist* (5, 6, 7) for one or more nuclear families in the income-sharing arrangement, since the availability of data was determined by eyeballing raw data files. The family relations implied

by codes 5, 6, and 7 are correct. An extended family has a value of missing data on the income and wealth variables in the INCEXT file if any family in the income-sharing arrangement had missing income data. Determination of missing data on income for the INCEXT file was done by a computer using the actual income files (INCOME10.SAS3) and did not depend on the information obtained by eyeballing the data.

HIGRDF Highest school grade attained by members of nuclear family.

Input Variables:

KID000 file: GRADOK
PAR000 file: GRADOM, GRADOP

Coding:

HIGRDF is the highest school grade attained by any nuclear family member living with the family at the time of the 425 Census.

MB0000 Number of male children of 02 who are less than 1 year old and currently living with the nuclear family.

Input Variables:

KID000 file: ESTADO, AGEMOS, SEXCEN, PERSON

Coding:

MB0000 is the sum of children in the nuclear family with PERSON = 03-21 or 31-40 for whom SEXCEN = 1, AGEMOS \leq 12 and ESTADO = 1.

Notes and Comments:

See Note 4 in Appendix A.

MB0106 Number of male children of 02 who are 1-6 years old and currently living with the nuclear family.

Input Variables:

KID000 file: ESTADO, AGEMOS, PERSON, SEXCEN

Coding:

MB0106 is the sum of children in the nuclear family with PERSON = 03-21 or 31-40 for whom SEXCEN = 1, $12 \leq$ AGEMOS \leq 83 and ESTADO = 1

Notes and Comments:

See Note 4 in Appendix A.

MB0710 Number of male children of 02 who are 7-10 years old and currently living with the nuclear family.

Input Variables:

KID000 file: ESTADO, AGEMOS, PERSON, SEXCEN

Coding:

MB0710 is the sum of children in the nuclear family with PERSON = 03-21 or 31-40 for whom SEXCEN = 1, $84 \leq$ AGEMOS \leq 131 and ESTADO = 1.

Notes and Comments:

See Note 4 in Appendix A.

MB1114 Number of male children of 02 who are 11-14 years old and currently living with the nuclear family.

Input Variables:

KID000 file: ESTADO, AGEMOS, PERSON, SEXCEN

Coding:

MB1114 is the sum of children in the nuclear family with PERSON = 03-21 or 31-40 for whom SEXCEN = 1, $132 \leq$ AGEMOS \leq 179 and ESTADO = 1.

Notes and Comments:

See Note 4 in Appendix A.

MB1545 Number of male children of 02 who are 15-45 years old and currently living with the nuclear family.

Input Variables:

KID000 file: ESTADO, AGEMOS, PERSON, SEXCEN

Coding:

MB1545 is the sum of children in the nuclear family with PERSON = 03-21 or 31-40 for whom SEXCEN = 1, $180 \leq$ AGEMOS \leq 551 and ESTADO = 1.

Notes and Comments:

See Note 4 in Appendix A.

MB4659 Number of female children of 02 who are 46-59 years old and currently living with the nuclear family.

Input Variables:

KID000 file: ESTADO, AGEMOS, PERSON, SEXCEN

Coding:

MB4659 is the sum of children in the nuclear family with PERSON = 03-21 or 31-40 for whom SEXCEN = 1, 552 ≤ AGEMOS < 719, and ESTADO = 1.

Notes and Comments:

See Note 4 in Appendix A.

MB60UP Number of male children of 02 who are 60 years old or older and currently living with the nuclear family.

Input Variables:

KID000 file: ESTADO, AGEMOS, PERSON, SEXCEN

Coding:

MB60UP is the sum of children in the nuclear family with PERSON = 03-21 or 31-40 for whom SEXCEN = 1, AGEMOS ≥ 720 and ESTADO = 1.

Notes and Comments:

See Note 4 in Appendix A.

MOMPRS Index of presence of person 02 in nuclear family.

Input Variables:

PAR000 file: NUCFAM, PERSON, ESTADO

Coding:

- 1 If ESTADO = 1 for person number 02
- 1 Otherwise (i.e., ESTADO = 2 -5)
 - Census card for 02 missing from file. See Note 3, Appendix A.

MONTHC Month the 425 Census was taken.

Input Variables:

425 Census, 100 series cards, columns 18-19.

Coding:

MONTHC is the number of the month that the 425 Census was taken, ranging from 01-12.

Notes and Comments:

In order to obtain complete date of census, MONTHC must be used in conjunction with DAYCEN and YEARCE.

MT0000

Number of related males who are less than 1 year old and currently living with the nuclear family

Input Variables:

KID000 file: ESTADO, AGEMOS, PERSON, SEXCEN
PAR000 file: ESTADO, AGEMOS, PERSON, SEXCEN

Coding:

MT0000 is the sum of nuclear family members with PERSON = 01-90 for whom SEXCEN = 1, AGEMOS \leq 12 and ESTADO = 1.

Notes and Comments:

See Note 5 in Appendix A.

MT0106

Number of related males who are 1-6 years old and currently living with the nuclear family.

Input Variables:

KID000 file: ESTADO, AGEMOS, PERSON, SEXCEN
PAR000 file: ESTADO, AGEMOS, PERSON, SEXCEN

Coding:

MT0106 is the sum of nuclear family members with PERSON = 01-90 for whom SEXCEN = 1, $12 \leq$ AGEMOS \leq 83, and ESTADO = 1.

Notes and Comments:

See Note 5 in Appendix A.

MT0710

Number of related males who are 7-10 years old and currently living with the nuclear family.

Input Variables:

KID000 file: ESTADO, AGEMOS, PERSON, SEXCEN
PAR000 file: ESTADO, AGEMOS, PERSON, SEXCEN

Coding:

MT0710 is the sum of nuclear family members with PERSON = 01-90 for whom SEXCEN = 1, $84 \leq \text{AGEMOS} \leq 131$ and ESTADO = 1.

Notes and Comments:

See Note 5 in Appendix A.

MT1114 Number of related males who are 11-14 years old and currently living with the nuclear family.

Input Variables:

KID000 file: ESTADO, AGEMOS, PERSON, SEXCEN
PAR000 file: ESTADO, AGEMOS, PERSON, SEXCEN

Coding:

MT1114 is the sum of nuclear family members with PERSON = 01-90 for whom SEXCEN = 1, $132 \leq \text{AGEMOS} \leq 179$ and ESTADO = 1.

Notes and Comments:

See Note 5 in Appendix A.

MT1545 Number of related males who are 15-45 years old and currently living with the nuclear family.

Input Variables:

KID000 file: ESTADO, AGEMOS, PERSON, SEXCEN
PAR000 file: ESTADO, AGEMOS, PERSON, SEXCEN

Coding:

MT1545 is the sum of nuclear family members with PERSON = 01-90 for whom SEXCEN = 1, $180 \leq \text{AGEMOS} \leq 551$ and ESTADO = 1.

Notes and Comments:

See Note 5 in Appendix A.

MT4659 Number of related males who are 46-59 years old and currently living with the nuclear family.

Input Variables:

KID000 file: ESTADO, AGEMOS, PERSON, SEXCEN
PAR000 file: ESTADO, AGEMOS, PERSON, SEXCEN

Coding:

MT4659 is the sum of nuclear family members with PERSON = 01-90 for whom SEXCEN = 1, $552 \leq \text{AGEMOS} \leq 719$ and ESTADO = 1.

Notes and Comments:

See Note 5 in Appendix A.

MT60UP

Number of related males who are 60 years and older and currently living with the nuclear family.

Input Variables:

KID000 file: ESTADO, AGEMOS, PERSON, SEXCEN

PAR000 file: ESTADO, AGEMOS, PERSON, SEXCEN

Coding:

MT60UP is the sum of nuclear family members with PERSON = 01-90 for whom SEXCEN = 1, $\text{AGEMOS} \geq 720$ and ESTADO = 1.

Notes and Comments:

See Note 5 in Appendix A.

NEMBRZ

Total number of pregnancies experienced by 02.

Input Variables:

425 Census, card 000, columns 31-32

Coding:

NEMBRZ is the number of pregnancies experienced by 02. Pregnancy includes stillbirths and spontaneous and induced abortions.

Notes and Comments:

NEMBRZ is coded missing when the question was not answered or when the family had been included with another nuclear family. See Note 6 in Appendix A.

NHIJOS

Total number of living children of 02.

Input Variables:

425 Census, card 000, columns 33-34

Coding:

NHIJOS is the number of living children of 02.

Notes and Comments:

NHIJOS is coded missing when the question was not answered or when the nuclear family has been included with another nuclear family. See Note 6 in Appendix A.

NMORTI Total number of stillbirths experienced by 02.

Input Variables:

425 Census, card 000, column 37

Coding:

NMORTI is the number of stillbirths experienced by 02.

Notes and Comments:

NMORTI is coded missing when the question was not answered or when the nuclear family has been included with another nuclear family. See Note 6 in Appendix A.

NMUERT Total number of children born alive to 02 who have died subsequently.

Input Variables:

425 Census, card 000, column 35-36

Coding:

NMUERT is the number of 02's children who died subsequent to birth.

Notes and Comments:

NMUERT is coded missing when the question was not answered or when the nuclear family has been included with another nuclear family. See Note 6 in Appendix A.

NPERFM Reported number of people in nuclear family.

Input Variables:

425 Census, card 000, columns 38-39

Coding:

NPERFM is the number of people in the nuclear family as reported by respondent at time of 425 Census.

Notes and Comments:

NPERFM is coded missing when the question was not answered

or when the nuclear family has been included with another nuclear family. See Note 6 in Appendix A.

The value of NPERFM does not always equal the value of CNTPIF since respondents' reports often include those who had an ESTADO other than 1 or those with a person ID number greater than 90. CNTPIF is a tally of only those people with PERSON = 01-90 and ESTADO = 1.

NPERVI Reported number of people living in house.

Input Variables:

425 Census, card 000, columns 40-41

Coding:

NPERVI is the number of people living in the house as reported by respondent at time of 425 Census.

Notes and Comments:

NPERVI is coded missing when the question was not answered or when the nuclear family has been included with another nuclear family. See Note 6 in Appendix A.

NUCFAM Nuclear family ID number.

Input Variables:

425 Census, 100 series cards, columns 9-11.

Coding:

NUCFAM is the nuclear family ID number assigned to the family at the beginning of the study.

Notes and Comments:

In order to obtain an ID which distinguishes a nuclear family from all other nuclear families, this variable must be used in conjunction with VILAGE.

NUM006 Number of children aged 0 to 6 years in nuclear family.

Input Variables:

KID000 file: AGEMOS, ESTADO

Coding:

NUM006 is the sum of children in the nuclear family for whom PERSON = 03-90, $0 < \text{AGEMOS} < 83$ and ESTADO = 1.

Notes and Comments:

Count includes only those for whom age in months could be calculated. See AGEMSF to determine which nuclear families have members with missing birth date data.

NUM714 Number of children aged 7-14 years in nuclear family.

Input Variables:

KID000 file: AGEMOS, ESTADO

Coding:

NUM714 is the sum of children in the nuclear family for whom PERSON = 03-90, 84 ≤ AGEMOS ≤ 179 and ESTADO = 1.

Notes and Comments:

Count includes only those for whom age in months could be calculated. See AGEMSF to determine which nuclear families have members with missing birth date data.

NUMBFA Number of nuclear families in the extended family.

Input Variables:

CENNUC file: EXTFAM, NUCFAM

Coding:

Actual number of nuclear families ranging from 1-4.

NURSNG Mother's current nursing status.

Input Variables:

425 Census, card 000, column 43

Coding:

0 is not currently nursing
1 says is currently nursing but it was not confirmed
2 is currently nursing (it was confirmed)
° missing data (see also Note 6 in Appendix A)

OCCA01 Current occupation of 01.

Input Variables:

PAR000 file: OCACTP

Coding:

See OCACTK in KID000 file.

OCCA02 Current occupation of 02.

Input Variables:

PAR000 file: OCACTM

Coding:

See OCACTK in KID000 file.

OCCP01 Usual occupation of 01.

Input Variables:

PAR000 file: OCPRNP

Coding:

See OCACTK in KID000 file.

OCCP02 Usual occupation of 02.

Input Variables:

PAR000 file: OCPRNM

Coding:

See OCACTK in KID000 file.

PARPRS Index of presence of both parents in nuclear family.

Input Variables:

PAR000 file: PERSON, ESTADO

Coding:

- 1 If ESTADO = 1 for both person numbers 01 and 02
- 1 Otherwise (ESTADO = 2 - 5)
 - ° Census card for father and mother missing from file.
See Note 3 in Appendix A.

PRGNOW Mother's current pregnancy status.

Input Variables:

425 Census, card 000, column 42

Coding:

- 0 is not currently pregnant
- 1 says she is pregnant but not confirmed
- 2 is pregnant (confirmed)
- ° missing data

Notes and Comments:

See also Note 6 in Appendix A.

RELATN Relation of head nuclear of family to head of extended family.

Input Variables:

425 Census, card 000, columns 44-45

Coding:

See code for RESPDT below for codes.

Notes and Comments:

See Note 7 in Appendix A.

RELSNR Religion of 01.

Input Variables:

425 Census, card 000, column 30

Coding:

- 1 Catholic
- 2 Protestant
- 3 Other

RELSRA Religion of 02.

Input Variables:

425 Census, card 000, column 29

Coding:

- 1 Catholic
- 2 Protestant
- 3 Other

RESPDT Identification number of respondent

Input Variables:

425 Census, card 000, columns 12-13

Coding:

- 01 head (father) of nuclear family
- 02 wife of head (mother)
- 03-20 child of 01 and of 02
- 21-30 child of 01 but not of current 02
- 31-40 child of 02 but not of current 01
- 41-50 niece or nephew of 01
- 51-60 niece or nephew of 02
- 61-70 grandchild of 01 or 02
- 71-80 adopted child
- 81 father of 01
- 82 mother of 01
- 83 father of 02
- 84 mother of 02
- 85-90 other relative of 01 or 02
- 91-98 nonrelated person

SEXMSF Index of missing data on sex of any nuclear family member.

Input Variables:

PAR000 file: ESTADO, SEXCEN
KID000 file: ESTADO, SEXCEN

Coding:

SEXMSF is a tally of people in the nuclear family at the time of the 425 Census (ESTADO = 1) with missing data on sex. A missing data code was assigned when a nuclear family had been given an address but members were included in the 425 Census of another family. See Note 1 in Appendix A.

SEXRES Sex of respondent specified in RESPDT.

Input Variables:

425 Census, card 000, column 14

Coding:

1 male
2 female

VILAGE Identification number of village in which nuclear family resides.

Input Variables:

425 Census, card 000, columns 7-8

Coding:

03 Santo Domingo
06 Conacaste
08 Espiritu Santo
14 San Juan
31 Colonias San Antonio and El Cortijo
32 San Miguel Petapa

YEARCE Year in which 425 Census was taken.

Input Variables:

425 Census, card 000, columns 20-21

Coding:

YEARCE is the last 2 digits of the year in which the 425 Census was taken, ranging from 75-76.

Notes and Comments:

In most cases the year is 1975. Some nuclear families apparently were missed in 1975. They were discovered during the cleaning of R10 in 1976 since they had income data but no 425 Census data. The 425 Census was done for these families at this later date, and explains why some have a 1976 date for the Census.

In order to obtain complete date of Census, YEARCE must be used in conjunction with DAYCEN and MONTHC.

IV. DEMOGRAPHIC VARIABLES AT THE EXTENDED FAMILY LEVEL

4.1 GENERAL COMMENTS ON THE CENEXT FILE

The CENEXT file contains demographic data for the *extended families* composed of the nuclear families who appear in the 425 Census. Each observation in the CENEXT file corresponds to an extended family. Here an extended family includes all the nuclear families who share household economic resources including means of production as well as income. There may be one or more nuclear families in an extended family. If there are more than one, they may or may not all be living in the same house. If there is only one family in the extended family, the value of many of the variables in this file will be the same as the value of the corresponding variables for this family in the CENNUC file. More detailed information about the definition of extended families and the identification of such families is given in the description of HHFORM in this file.

Most of the variables contained in the file were constructed from the CENNUC file based originally on the 425 Census. In many cases more detail about a variable is given in the description of the component variable in the CENNUC file.

4.2 DESCRIPTION OF VARIABLES CONTAINED IN CENEXT FILE.

AGEMIX: Index of missing data on age of any extended family member.

Input Variables:

CENNUC file: AGEMIS

Coding:

0 if no extended family member has missing birth date data
1 if one or more extended family members have missing birth date data.

Notes and Comments:

This variable indicates whether the following household composition variables include *all* extended family members: DR006X, DR714X, EACONX, NM006X, NM714X. Since occasionally birth dates were not known, those extended families for whom AGEMIX = 1 will exclude one or more people from the count. Generally, missing data appears for grandparents and/or more distant relatives (e.g., cousins, nieces, uncles, etc.) who live in the extended family.

AGEMSX: Index of missing birth date data for any extended family member.

Input Variables:

CENNUC file: AGEMSF

Coding:

AGEMSX is a tally of people in the extended family with missing data on birth date.

CENPIV: Number of people in house as reported on 425 Census.

Input Variables:

CENNUC file: NPERVI

Coding:

CENPIV is the highest value of NPERVI on the CENNUC file reported by each of the nuclear families in the extended family. The variable is set to missing if the extended family is composed of one nuclear family and that family had no 000 card on the 425 Census.

Notes and Comments:

See Note 1 in Appendix B.

CENPIX: Tally of reported number of people in the extended family.

Input Variables:

CENNUC file: NPERFM

Coding:

CENPIX is the value of

$$\sum_{i=1}^n \text{NPERFM}_i$$

where NPERFM_i = NPERFM from CENNUC file for nuclear family i ($i = 1, 2, \dots, n$)

n = number of nuclear families in the extended family

The variable is set to missing if the extended family is composed of one nuclear family and that family had no 000 series card on the 425 Census. See Note 1 in Appendix B.

CNTFIX: Tally of nuclear families in extended family.

Input Variables:

CENNUC file: NUMBFA

Coding:

Actual member of nuclear families ranging from 1-4.

CNTPIX Number of related people with complete birth date data currently living in the extended family.

Input Variables:

CENNUC file: CNTPRF

Coding:

CNTPIX is the value of

$$\sum_{i=1}^n \text{CNTPRF}_i$$

where CNTPRF_i = CNTPRF from CENNUC file for nuclear family i ($i = 1, 2, \dots, n$)

n = number of nuclear families in the extended family

Notes and Comments:

All people with ID numbers 01-90 were counted; however, unlike CNT2PX only those for whom age in months could be calculated were included in the count. See AGEMX to determine which extended families have members with missing birth date data.

CNT2PX: Number of related people currently living with extended family.

Input Variables:

CENNUC file: CNTPIF

Coding:

CNT2PX is the value of

$$\sum_{i=1}^n \text{CNTPIF}_i$$

where CNTPIF_i = CNTPIF from CENNUC file for nuclear family i ($i = 1, 2, \dots, n$)

n = number of nuclear families in the extended family

Notes and Comments:

CNT2PX differs from CNT1PX in that all related people (PERSON = 01-90) were counted regardless of whether or not age in months could be calculated for them, whereas in CNT1PX, only those with complete birth date data were counted. See also Note 1 in Appendix B.

DENOMX: Number of adults aged 15 to 60 years in extended family.

Input Variables:

CENNUC file: DENOMR

Coding:

DENOMX is the value of

$$\sum_{i=1}^n \text{DENOMR}_i$$

where $DENOMR_i = DENOMR$ for nuclear family i
($i = 1, 2, \dots, n$)

$n =$ number of nuclear families in the
extended family.

Notes and Comments:

Count includes only those for whom age in months could be calculated. See AGEMSX to determine which extended families have members with missing age data.

DR006X: Number of children 0-6 years old per adult in the extended family.

Input Variables:

CENEXT file: DR006X DENOMX

Coding:

DR006X is the value of
 $(NM006X/DENOMX) * 100$

Notes and Comments:

The components NM006X and DENOMX include only those individuals for whom age in months could be calculated. See AGEMSX to determine which extended families have members with missing age data.

DR714X: Number of children 7-14 years old per adult in the extended family.

Input Variables:

CENEXT file: NM714X, DENOMX

Coding:

DR714X is the value of
 $(NM714X/DENOMX) * 100$

Notes and Comments:

Components NM714X and DENOMX include only those individuals for whom age data were available. See AGEMSX to determine which extended families have members with missing birth date data.

EACONX: Number of equivalent adult consumers in the extended family.

Input Variables:

CENNUC file: EACONF

Coding:

EACONX is the value of

$$\frac{1}{n} \left(\sum_{i=1}^n \text{EACONF}_i \right) * 100$$

where EACONF_i = EACONF for nuclear family i
($i = 1, 2, \dots, n$)

n = number of nuclear families in the
extended family

Notes and Comments:

See EACONF for details about construction of variable.
EACONX includes only those individuals for whom age in
months could be calculated. See MSEACX to determine which
extended families have members with missing birth date data.

EXTFAM: Identification numbers of extended family.

Input Variables:

CENNUC file: EXTFAM

Coding:

See EXTFAM on CENNUC file.

Notes and Comments:

See EXTFAM on CENNUC file.

HHMODE: Index of type of household.

Input Variables:

CENNUC file: HHFORM

Coding:

See HHFORM on CENNUC file.

HIGRDX: Highest school grade attained by members of extended family.

Input Variables:

CENNUC file: HIGRDF

Coding:

HIGRDX is the maximum value of grades attained (HIGRDF) for all members of the extended family who were living with the extended family at the time of the interview (e.g., ESTADO = 1)

NM006X: Number of children aged 0 to 6 years in extended family.

Input Variables:

CENNUC file: NUM006

Coding:

NM006X is the value of

$$\sum_{i=1}^n \text{NUM006}_i$$

where $\text{NUM006}_i = \text{NUM006}$ for nuclear family i
($i = 1, 2, \dots, n$)

$n =$ number of nuclear families in the extended family

Notes and Comments:

Count includes only those for whom age in months could be calculated. See AGEMSX to determine which extended families have members with missing birth date data.

NM714X: Number of children aged 7 to 14 years in the extended family.

Input Variables:

CENNUC file: NUM714

Coding:

NM714X is the value of

$$\sum_{i=1}^n \text{NUM714}_i$$

where $\text{NUM714}_i = \text{NUM714}$ for nuclear family i
($i = 1, 2, \dots, n$)

$n =$ number of nuclear families in the extended family

Notes and Comments:

Count includes only those for whom age in months could be calculated. See AGEMSX to determine which extended families have members with missing birth date data.

SEXMSX: Index of missing data on sex of any extended family member.

Input Variables:

CENNUC file: SEXMSF

Coding:

SEXMSX is the value of

$$\sum_{i=1}^n \text{SEXMSF}_i$$

where $\text{SEXMSF}_i = \text{SEXMSF}$ for nuclear family i
($i = 1, 2, \dots, n$)

n = number of nuclear families in the
extended family

Notes and Comments:

In all cases, the value of this variable is 0, implying that no extended family had missing data on the sex of any of its members.

VILAGE: Identification numbers of Village in which extended family resides.

Input Variables:

CENNUC file: VILAGE

Coding:

03 Santo Domingo

06 Conacaste

08 Espiritu Santo

14 San Juan

31 Colonia San Antonio and El Cortijo

32 San Miguel Petapa

V. INCOME AND WEALTH AND SOCIOECONOMIC VARIABLES AT THE NUCLEAR FAMILY LEVEL

5.1 GENERAL COMMENTS ON THE INCOME10 FILE

The INCOME10 file contains income and wealth data for the *nuclear families* in the study communities. Each observation in the INCOME10 file corresponds to a nuclear family who was administered an R10 questionnaire. If a nuclear family appears in the CENNUC file but not in the INCOME10 file, it is because there was no completed R10 questionnaire for that nuclear family.

Some of the variables appearing in this file were taken from cards 000 and 001 of the 425 Census and pertain to the socioeconomic and demographic characteristics of the nuclear family. Cleaner versions of the variables based on the 425 Census appear in the CENNUC file.

Many of the income and wealth variables are measured in the Guatemalan monetary unit, the Quetzal. One quetzal is equivalent to one U.S. dollar.

Some income variables are extremely low and/or negative. The original questionnaires for these cases have been checked for coding errors. Those who continued to have low or negative incomes during 1974: a) may have actually gone into debt that year (e.g. had a bad harvest due to drought, sold crops at a time when prices were low and taken a loss, etc.), b) may have been receiving regular transfers of food, clothing, etc., from others or may have had other sources of income which were not adequately captured by the R10 questionnaire, c) may have underestimated income from the various sources or d) may have overestimated costs. This could be for the following reasons: a) fear that the interviewers represented the government, b) recall error, or c) imperfect information due to the nature of income in these settings. (Income in-kind, home consumption of agricultural production and household enterprises characterized by very informal systems of accounting cash flows often make it impossible for individuals to give precise income information).

Cases of negative and very low income have been left on the files, enabling the researcher to decide how to treat them. Likewise, high values have been checked for coding errors. These values, too have been left on the files.

5.2 DESCRIPTION OF VARIABLES CONTAINED IN INCOME10 FILE

CINRST Total 1974 interest costs in quetzales.

Input Variables:

R10 survey, card 001, columns 63-64, 65-67

Coding:

CINRST is the product of monthly interest payments made (columns 63-64) times the number of months such payments were made (columns 65-67) if the number of months \leq 12. Otherwise, CINRST is the product of the monthly interest payments times 12.

CIVMAD Current living status of 02.

Input Variables:

425 Census, 100 series cards, column 46.

Coding:

- 1 Present in household at the time of the 425 Census interview. Also includes women who are away from the community, but maintain direct links with the family by visiting regularly and continuing to be responsible for the family (e.g., works in Guatemala City during the week and returns on weekends to village). This does not include those who live in another country or have established another household elsewhere.
- 2 Dead
- 3 Lives in house but is considered part of another nuclear family.
- 4 Lives away from house but in the community and forms part of another nuclear family.
- 5 Lives outside the community independent of the nuclear family.

CIVPAD Current living status of 01

Input Variables:

425 Census, 100 series cards, column 46.

Coding:

See CIVMAD for code.

CIVSTA Civil status head of household.

Input Variables:

425 Census, card 000, column 27.

Coding:

- 1 Single without children
- 2 Single without formal union
- 3 Single mother
- 4 Mother without stable union
- 5 Consensual union
- 6 Married
- 7 Separated or divorced
- 8 Widow

CLABOR Costs in quetzales of hired labor used in production of all crops grown in 1974.

Input Variables:

R10 survey, 500 series cards, columns 27-28, 29-30, 31-32, 33, 34-35, 36-37, 38-39, 40-41, 42, 43-44, 45-46, 47, 48-49, 50-51, 52, 53-54, 55-56, 57, 58-59, 60-61, 62-63, 64-65, 66-67, 68-69.

Coding:

CLABOR is the sum of days worked by others (e.g., those coded "96" in columns 27-28) multiplied by the price of hired labor for that village.

Formula used to calculate amount of hired labor:

$$\begin{aligned} & \Sigma \left[\left(\frac{\text{col } 29-30}{\text{col } 33} \right) + \left(\frac{\text{col } 31-32}{\text{col } 33} \right) + (\text{col } 34-35) + (\text{col } 36-37) + \left(\frac{\text{col } 38-39}{\text{col } 42} \right) \right. \\ & + \left(\frac{\text{col } 40-41}{\text{col } 42} \right) + \left(\frac{\text{col } 43-44}{\text{col } 47} \right) + \left(\frac{\text{col } 45-46}{\text{col } 47} \right) + \left(\frac{\text{col } 48-49}{\text{col } 52} \right) + \left(\frac{\text{col } 50-51}{\text{col } 52} \right) \\ & + \left(\frac{\text{col } 53-54}{\text{col } 57} \right) + \left(\frac{\text{col } 55-56}{\text{col } 57} \right) + (\text{col } 58-59) + (\text{col } 60-61) + (\text{col } 62-63) \\ & \left. + (\text{col } 64-65) + (\text{col } 66-67) + (\text{col } 68-69) \right] \end{aligned}$$

where "col" refers to column on 500 series card. If a blank appeared in columns 29-30, 33, 42, 47, 52 or 57, they were assumed to equal "1". If a 2 or 3 appeared, they were assumed to equal "1/2". A "tajea" is defined as a "days worth of work" by the rural farmers.

The following village wage rates based on R10b (Community Price Questionnaire) were used.

Village No.	Wage Rate (quetzales)
03	.75
06	.85
08	.90
14	.85
31-32	1.05

CLANDP Rental costs in quetzales of land used in 1974.

Input Variables:

R10 survey, 100 series cards, columns 45-47, 34 and 43.

Coding:

CLANDP is the sum of rent paid on each parcel of land (columns 45-47) for which respondent indicated that a) someone outside the nuclear or extended family owned it (e.g., column 34 = 3 or 4) and b) rent was paid (e.g., column 43 = 1 or 2).

CLSMAD Is 02 attending any special classes?

Input Variables:

425 Census, 100 series cards, column 76.

Coding:

0 No
1 Yes

CONDUR Index of ownership of consumer durables.

Input Variables:

R10 survey, 900 series cards, columns 22-23, 24-25, 26.

Coding:

CONDUR is the weighted sum of the following durables owned by the nuclear family. Weights are shown below. Durables in good condition received slightly more weight than those in bad condition. Data on condition were obtained from the respondent himself who simply was asked if it was in good or bad condition.

<u>Consumer Durable</u>	<u>Good Condition</u>	<u>Bad Condition</u>
Radio	2	1
T.V.	6	5
Sewing Machine	4	3
Refrigerator	6	5

In determining weights, cost, type of energy used, and distribution of durables among households was considered. It is assumed that sewing machines, refrigerators and T.V.s are consumer durables. In a context where home enterprises are so common, none of these need be strictly a consumer durable but may serve a productive function as well. For example, the refrigerator may hold cold pop for sale; the T.V. may be used to attract people to a small family owned restaurant.

CONSUP

Variable used for file construction of schooling behavior for those families with children aged 7-14 in 1975.

Input Variables:

R10, R08 surveys.

Coding:

- 1 & 2 Nuclear family has R08 data but no income data
- 3 & 4 Nuclear family has R08 and income data
- 5 - 7 Nuclear family has income data but no R08 data

CVARIN

Total cost in quetzales of variable inputs to agricultural crop production in 1974.

Input Variables:

Corrected version of the April 1977 CROPHEAD file, columns 31-38.

Coding:

CVARIN is the total cost in quetzales of regular and improved seed, fertilizer and insecticides for production

of the following 6 major crops: corn, beans, tomatoes, jalapeño chiles, serrano chiles and barley. Thus, CVARIN is the sum of columns 31-38 across all cards appearing in the CROPHEAD file for a nuclear family.

DAYMAD Day of birth of 02.

Input Variables:

425 Census, 100 series cards, columns 52-53.

Coding:

DAYMAD is the number of the day ranging from 01-31.

Notes and Comments:

In order to obtain exact birth date, this variable must be used in conjunction with MONMAD and YERMAD. DAYMAD appears on the PAR000 as DAYNAC for person number 02.

DAYPAD Day of birth of 01.

Input Variables:

425 Census, 100 series cards, columns 52-53.

Coding:

See DAYMAD above.

Notes and Comments:

In order to obtain exact birth date, this variable must be used in conjunction with MONPAD and YERPAD. DAYPAD appears on the PAR000 file as DAYNAC for person number 01.

DXHOME Index of housing quality.

Input Variables:

425 Census, card 000, columns 27-29

Coding:

DXHOME is a scale ranging from 3-16 (poor to good) indicating the quality of housing, as indicated by material of floor, roof and walls. It was calculated by summing columns 27, 28, and 29 (e.g., Σ MFLOOR + MTROOF + MTWALL).

DXUTLY Index of public utilities in house.

Input Variables:

425 Census, card 001, columns 32-35

Coding:

DXUTLY is a scale ranging from 3 to 12 (poor to good) indicating extent of public utilities (including electricity, sanitation, sewers and water) found in the house. It was calculated by summing columns 32, 33, 34, and 35 (e.g., Σ LIGHTS + SANITN + SEWDSP + SSWATR).

EXSTRC

Variable used for construction of file of schooling behavior for those nuclear families with children aged 7 to 14 in 1975.

Input Variables:

R08, R10.

Coding:

- 1 R08 but no R10 data available (code values 1 & 2 on CONSUP)
- 2 R08 and R10 data available (code values 3 & 4 on CONSUP)
- 3 No R08 but R10 data available (code values 5-7 on CONSUP)

GRAMAD

Grades completed by 02.

Input Variables:

425 Census, 100 series cards, columns 74-75.

Coding:

GRAMAD is the actual number of school grades completed by person number 02.

Notes and Comments:

GRAMAD appears in the PAR000 and CENNUC files as GRADOM and GRAD02 respectively.

GRAPAD

Grades completed by 01.

Input Variables:

425 Census, 100 series cards, columns 74-75.

Coding:

See GRAMAD above.

Notes and Comments:

GRAPAD appears in the PAR000 and CENNUC files as GRADOP and GRAD02 respectively.

KITCHEN Location of kitchen.

Input Variables:

425 Census, card 001, column 30

Coding:

- 1 In the sleeping area
- 2 In a separate place
- 3 Incorporated into the living area

LIGHTS Index for presence of electricity in house.

Input Variables:

425 Census, card 001, column 32.

Coding:

- 0 no electricity present
- 1 electricity present

LITMAD Literacy of 02.

Input Variables:

425 Census, 100 series cards, column 73.

Coding:

- 0 Does not know how to read and write
- 1 Reads and writes with difficulty
- 2 Reads and writes without difficulty

Notes and Comments:

LITMAD appears in the PAR000 and CENNUC files as ALFAMA and ALFA02 respectively.

LITPAD Literacy of 01.

Input Variables:

425 Census, 100 series cards, column 73.

Coding:

See LITMAD for code.

Notes and Comments:

LITPAD appears in the PAR000 and CENNUC files as ALFAPA and ALFA01 respectively.

MFLOOR Material of floor.

Input Variables:

425 Census, card 001, column 27

Coding:

- 1 dirt
- 2 clay or brick
- 3 cement
- 4 mosaic
- 5 wood

MIGMAD Type of place 02 migrated to.

Input Variables:

425 Census, 100 series cards, column 63.

Coding:

- 1 Migrated to plantation (finca)
- 2 Migrated to a small village (caserio) without a local government
- 3 Migrated to a village (aldea--smallest geographical division with a local government)
- 4 Migrated to a municipal capital or county seat (cabecera municipal)
- 5 Not applicable, mother did not migrate (e.g., ESTADO = 1-4)

MIGPAD Type of place 01 migrated to.

Input Variables:

425 Census, 100 series cards, column 63.

Coding:

See code for MIGMAD above.

MONMAD

Birth month of 02.

Input Variables:

425 Census, 100 series cards, columns 54-55.

Coding:

MONMAD is the number of the month in which the individual was born, ranging from 01 to 12.

Notes and Comments:

In order to obtain exact birth date, this variable must be used in conjunction with DAYMAD and YERMAD. MONMAD appears in the PAR000 file as MESNAC for person number 02.

MONPAD

Birth month of 01.

Input Variables:

425 Census, 100 series cards, columns 54-55.

Coding:

See MONMAD for detail.

Notes and Comments:

In order to obtain exact birth date, this variable must be used in conjunction with DAYPAD and YERPAD. MONPAD appears in the PAR000 file as MESNAC for person number 01.

MTROOF

Material of roof.

Input Variables:

425 Census, card 001, column 28

Coding:

- 1 straw or similar material
- 2 thatch
- 3 corrugated fiberglass or aluminum
- 4 tile

MTWALL Index of material that walls of house are made of.

Input Variables:

425 Census, card 001, column 29

Coding:

- 1 cane or similar material
- 2 simple woven cane
- 3 woven cane reinforced with mud or adobe plaster
- 4 simple adobe
- 5 adobe reinforced with straw
- 6 wood
- 7 tile or block

NODURB Count of number of durables owned by household as reported in 425 Census.

Input Variables:

425 Census, card 001, columns 36-43

Coding:

NODURB is the simple sum of durables owned by household. Durables include:

- (a) radio
- (b) record player
- (c) television
- (d) bicycle
- (e) motorcycle
- (f) automobile
- (g) sewing machine
- (h) refrigerator

The range is from 0 (owns none) to 8 (owns one or more of each item).

NOROOM Number of rooms in the house.

Input Variables:

425 Census, card 001, columns 25-26

Coding:

NOROOM is the total number of rooms in the house in which the nuclear family resides.

NUCFAM Identification number of nuclear family.

Input Variables:

425 Census, card 000, columns 9-11

Coding:

NUCFAM is the nuclear family ID number assigned to the family at the beginning of the study.

Notes and Comments:

In order to obtain an ID which distinguishes the family from all other nuclear families, this variable must be used in conjunction with VILAGE.

OCACTM Current occupation of 02.

Input Variables:

425 Census, 100 series cards, columns 79-80.

Coding:

See OCACTK on the KID000 file.

Notes and Comments:

Additional occupation information for women is available in the R11 and R09 surveys. Rural Guatemalan (and some urban) women tend not to think of themselves as employed in the labor market sense but may actually be contributing to family income by working in the household enterprise (e.g., housefront store) or on family farms. The approaches used in these other questionnaires may get at their employment more completely.

OCACTP Current occupation of 01.

Input Variables:

425 Census, 100 series cards, columns 79-80.

Coding:

See OCACTK on KID000 file for codes.

OCPRNM Principal or usual occupation of 02.

Input Variables:

425 Census, 100 series cards, columns 77-78.

Coding:

See OCACTK on KID000 file for codes.

OCPRNP Principal or usual occupation of 01.

Input Variables:

425 Census, 100 series cards, columns 77-78.

Coding:

See OCACTK on KID000 file for codes.

OWNHME Ownership of house.

Input Variables:

425 Census, card 001, column 24

Coding:

- 1 living with neighbors
- 2 domestic employee
- 3 living with relatives
- 4 renting
- 5 ceded
- 6 paying mortgage
- 7 own home
- 8 other

OWNPRO Does nuclear family own property around home?

Input Variables:

425 Census, card 001, column 22

Coding:

- 0 none
- 1 no
- 2 yes

PERFAM Reported number of people in nuclear family.

Input Variables:

425 Census, card 000, columns 38-39. -39.

Coding:

PERFAM is the number of people in nuclear family as reported by respondent at the time of the 425 Census.

Notes and Comments:

A cleaner version of PERFAM appears as NPERFM on the CENNUC file.

PERHSE Reported number of people living in house.

Input Variables:

425 Census, card 000, columns 40-41

Coding:

PERHSE is the number of people living in the house as reported by respondent at the time of the 425 Census.

Notes and Comments:

A cleaner version of PERHSE appears as NPERVI on the CENNUC file.

PRODUR Index of ownership of producer durables.

Input Variables:

R10 survey, 900 series cards, columns

Coding:

PRODUR is the weighted sum of the following producer durables mentioned as being owned by respondent. Weights for durables in good and bad condition are shown below. In estimating weights for producer durables, both potential contribution to production as well as cost were considered.

<u>Type of Producer Durable</u>	<u>Good Condition</u>	<u>Bad Condition</u>
Wooden hoe	2	1
Metal hoe	4	3
Sprinkler	4	3
Water pump (irrigation)	10	9

<u>Type of Producer Durable</u>	<u>Good Condition</u>	<u>Bad Condition</u>
Cask-barrel	.25	.20
Stable	12	11
Tractor	12	11
Wagon (hand)	6	5
Mill with motor	8	7
Degrainer	6	5
Pick-up truck	10	9
Work room	12	11

Notes and Comments:

The portion of the questionnaire dealing with ownership of producer durables was designed mainly to obtain information on the type of durables owned by rural subsistence farmers and does not measure very well the types of producer capital used in the semi-urban areas (VILAGE = 31 and 32). In an attempt to remedy this, considerable weight was given to "work room" since this applies to non-agricultural types of household production. However, since no distinction was made in the questionnaire, a fully equipped butcher shop, for example, would receive the same score as a simple cobbler shop. Therefore, this index or any other that might be created from the data may be poor reflections of semi-urban household productive capital.

PROMAD Type of place 02 came from.

Input Variables:

425 Census, 100 series cards, column 63.

Coding:

- 1 Originated in plantation (finca)
- 2 Originated in small village (caserio) political division so small that it does not have its own governing offices
- 3 Originated in village with local government offices (aldea)
- 4 Originated in municipal capital or county seat (cabecera municipal)
- 5 Not applicable, has migrated from village (e.g., ESTADO = 5) or has died (ESTADO = 2)

PROPAD Type of place 01 came from.

Input Variables:

425 Census, 100 series cards, column 63.

Coding:

See PROMAD above for codes.

SANIML Net income in quetzales from sale of animals in 1974.

Input Variables:

April 1975 version of VBL800 file, columns 6-10 and 11-15
(See Kasala, May 1979).

Coding:

SANIML is the difference between total income from the rental and sale of animals (columns 6-10) minus the costs of feed (columns 11-15). Costs do not include costs of feed for previous years, initial purchase price of animal or value of feed grown by family.

Notes and Comments:

Those households which actually raised animals as a means of producing income (and not as a way of holding wealth in a less liquid form) had income from sale of animals coded in YDERPR. This includes 2 cases: 32 565 who has a chicken farm and 31 216 who appears to raise animals as a means of income production. For case 32 014, who produces milk, costs given for animal production were included in calculation of net income from derived products, YDERPR.

SANITN Index of sanitary facilities available in house.

Input Variables:

425 Census, card 001, column 53

Coding:

- 1 none in house
- 2 latrine
- 3 septic tank
- 4 toilet and plumbing

SEWDSP Index of sewage facilities available in house.

Input Variables:

425 Census, card 001, column 34

Coding:

- 1 throw water in yard
- 2 covered drain
- 3 pipe network

SPADAD Total income from paid labor (wages and salary) received by 01 during 1974.

Input Variables:

R10 survey, 600 series cards, columns 23-24 and 37-40.

Coding:

SPADAD is the sum of the total pay for each wage paying activity (columns 37-40) if the person doing the activity (columns 23-24) had an ID number 01.

Notes and Comments:

This variable was taken from VBL600 file (columns 6-12) created by John Stein and Bernard Pillet from cleaned R10 questionnaire data.

SPAMOM Total income from paid labor (wages and salary) received by 02 during 1974.

Input Variables:

R10 survey, 600 series cards, columns 23-24 and 37-40.

Coding:

SPAMOM is the sum of the total pay for each wage paying activity (columns 37-40) if the person doing the activity (columns 23-24) had an ID number of 02.

Notes and Comments:

This variable was taken from VBL600 file (columns 13-19) created by John Stein and Bernard Pillet, from cleaned R10 questionnaire data.

SPAOT1 Total income from paid labor (wages and salary) received

by children of 01 or 02 during 1974.

Input Variables:

R10 survey, 600 series cards, columns 23-24 and 37-40.

Coding:

SPAOT1 is the sum of the total pay for each wage-paying activity (columns 37-40) if the people doing the activity (columns 23-24) had ID numbers of 03-40.

Notes and Comments:

This variable was taken from VBL600 file (columns 20-26) created by John Stein and Bernard Pillet from cleaned R10 questionnaire data.

SPAOT2 Total income from paid labor (wages and salary) performed by people other than 01, 02, or their children during 1974.

Input Variables:

R10 survey, 600 series cards, columns 23-24 and 37-40.

Coding:

SPAOT2 is the sum of the total pay for each wage-paying activity (columns 37-40) if the people doing the activity (columns 23-24) had ID numbers of 41-98.

Notes and Comments:

This variable was taken from VBL600 file (columns 27-33) created by John Stein and Bernard Pillet from cleaned R10 questionnaire data.

SSORCH Wage source of person number 01 in 1974.

Input Variables:

R10 survey, 600 series cards, columns 25-26.

Coding:

- 0 If head received no wages or salary during 1974
- 1 If largest source of household head's wage income was agricultural daywork "jornalero agricola"
- 2 If largest source of household head's wage income was non-agricultural daywork "jornalero no-agricola"
- 3 If household head received some wages but source was neither of above.

SSORCW Wage source of 02 in 1974.

Input Variables:

R10 survey, 600 series cards, columns 25-26.

Coding:

- 0 If wife had no wages or salary during 1974
- 1 If largest source of household wife's wage income was agricultural daywork "jornalero agricola"
- 2 If largest source of household wife's wage income was non-agricultural daywork "jornalero no-agricola"
- 3 If wife received some wages but source was neither of above

SSWATR Index of source of water to house.

Input Variables:

425 Census, card 001, column 35

Coding:

- 1 river, lake or public water hole
- 2 public well or spigot
- 3 well in the house
- 4 hooked up to public network

TRADUR Index of transportation durables owned by nuclear family in 1974.

Input Variables:

R10 survey, 900 level cards, columns 22-23, 24-25, 26.

Coding:

The value is the weighted sum of the following transportation durables owned by the nuclear family. Weights are shown below. Durables in good condition received slightly more weight than those in bad condition. Scaling was based on potential productivity and quetzal value relative to other durables in same class.

<u>Durable</u>	<u>Bad Condition</u>	<u>Good Condition</u>
Car	7	8
Motorcycle	4	5
Pick-up truck	7	8
Bicycle	1	2

TRNREG Income in quetzales from regular transfers to the nuclear family in 1974.

Input Variables:

R10 survey, card 001, columns 30-33, and 34-37

Coding:

TRNREG is the sum of income from pension or retirement payments (columns 30-33) plus income from family members or others living away from the household (columns 34-37).

TRNONE Income in quetzales from once and for all transfers to the nuclear family in 1974.

Input Variables:

R10 survey, card 001, columns 22-25 and 26-29

Coding:

TRNONE is the sum of transfers from accident or sickness benefits (columns 22-25) plus severance pay (columns 26-29).

TRNTOT Income in quetzales from all transfers to the nuclear family in 1974.

Input Variables:

INCOME10 file: TRNREG, TRNONE

Coding:

TRNTOT is the sum of once and for all transfers (TRNREG) and regular transfers to the family (TRNTOT).

VILAGE Village identification number.

Input Variables:

R10 survey, all cards, columns 7-8.

Coding:

03 Santo Domingo (rural)
06 Conacaste (rural)
08 Espiritu Santo (rural)
14 San Juan (rural)
31 Colonias San Antonio and El Cortijo
32 San Miguel Petapa

VLANDN Value of quetzales of land owned by nuclear family in 1974.

Input Variables:

April 1977 version of the VBLAND file, columns 14-21, 22-29, 30-37, 38-45 (See Kasala, May 1979).

Coding:

In order to construct the index, each parcel of land owned by the nuclear family was rated as being prime (flat, lowland and irrigated), good (flat, lowland no irrigation), mediocre (hill or slope), or unidentified (not planted). The total area of each type of land owned was calculated and each area then multiplied by the quetzal value of that type of land for the village in question. These four values were then summed. Thus VLANDN is the value of

$$\Sigma(\text{PRIME}) * P_{p,i} + (\text{GOOD})P_{g,i} + (\text{SOSO})P_{s,i} + (\text{ASABER})P_{a,i}$$

where PRIME = Area of prime cultivated land owned by nuclear family (columns 14-21)

GOOD = Area of good cultivated land owned by nuclear family (columns 22-29)

SOSO = Area of mediocre cultivated land owned by nuclear family (columns 30-37)

ASABER = Area of unidentified quality land owned by nuclear family (columns 38-45)

$P_{p,i}$ = Price of prime land in village i
(i = 03, ..., 32)

$P_{g,i}$ = Price of good land in village i
(i = 03, ..., 32)

$P_{s,i}$ = Price of mediocre land in village i
(i = 03, ..., 32)

$P_{a,i}$ = Price of unidentified land in village i
(i = 03, ..., 32)

Prices of each type of land were obtained from the R10b community price survey administered about the same time as the R10 survey. See Table below for village land prices used.

Notes and Comments:

On the R10 questionnaire form, land that was not planted was not given a quality rating. In cases where quality of land was unknown because it was not currently under cultivation, the following assumptions were made. If the plot was less than 5 cuerdas (one cuerda = 625 varas, vara = 2 ft. 9 in.) in area, it was given the price of good land. If it was 5-32 cuerdas in area, it was given the price of

Table 4

VILLAGE PRICES OF VARIOUS LAND TYPES

Village	Type of Land		
	Irrigated	Flat Plane	Hill or Slope
03	Q 313	Q 215	Q 125
06	Q 583	Q 400	Q 183
08	Q 932.5	Q 460	Q 375
14	Q 324	Q 222	Q 133
31 & 32	Q 2833	Q 2166	Q 1066

mediocre land. If it was greater than 32 cuerdas in area, it was valued at one-half the price of mediocre land. These assumptions were made because it was assumed that small plots of land not currently under cultivation were laying fallow for a short period of time. It is generally the case that subsistence farmers cannot leave large quantities of good land out of cultivation at any one time. Therefore, if larger plots were not being cultivated, it was assumed that it was because they were of poor quality land.

Land area measurements originally were given in "cuerdas" of 12, 16, and 20. All were converted to cuerdas of 16 for these calculations. A cuerda of 12 = 625 square varas, a cuerda of 16 = 1024 square varas, a cuerda of 20 = 1600 square varas where one square vara \approx one square yard.

VLANDX Value in quetzales of land owned by the extended family in 1974.

Input Variables:

April 1977 version of the VBLLAND file columns 46-53, 54-61, 62-69, and 70-77.

Coding:

The respondent was asked about land owned by the extended family as well as by his own nuclear family. This variable was constructed in a manner similar to VLANDN, however, land areas reported for *extended* family were used. (e.g., columns 46-53, 54-61, 62-69, and 70-77 for prime,

good, mediocre and unidentified land quality respectively). See VLANDN for details on variable construction and Notes and Comments.

WRKWIF Number of days person number 02 worked in agricultural activity during 1974.

Input Variables:

R10 survey, 500 series cards, columns 29-30, 31-32, 34-35, 36-37, 38-39, 40-41, 43-44, 45-46, 48-49, 50-51, 53-54, 55-56, 58-59, 60-61, 62-63, 64-65, 66-67, and 68-69.

Coding:

WRKWIF is the sum of the above columns.

Notes and Comments:

A "tarjea" is defined as a "days worth of work" by the rural farmer.

YAGPRO Net income in quetzales from agricultural production in 1974.

Input Variables:

INCOME10 file: YSCROP, YHMCON, CVARIN, CLABOR and YLANDR.

Coding:

YAGPRO is the sum of income from sale of crops (YSCROP) plus value of home consumption (YHMCON) minus the costs of a) variable inputs (CVARIN) b) hired labor (CLABOR) and c) land rental (YLANDR).

YDERPR Net income in quetzales from derived products 1974.

Input Variables:

R10 survey, 700 series cards, columns 24-27.

Coding:

YDERPR is the difference between net income from production of all derived products (columns 24-27) and some proportion of extraordinary costs (columns 40-42) to account for depreciation associated with capital used in production of these derived products. Depreciation was set at 20% for small capital items and 10% for large capital costs. This measure does not include home consumption of derived products, since such data were not collected.

Notes and Comments:

This variable was created using SUM7 and EXTC from columns 6-13 and 14-21 respectively of VBL700 file created by John Stein and Bernard Pillet. As a measure of net income for production of any particular product, they took "the smaller of i) reported net income and ii) reported gross income less ordinary expenses.

YERMAD Birth month of 02.

Input Variables:

425 Census, 100 series cards, columns 56-57.

Coding:

YERMAD is the two digits of the year in which the woman was born.

Notes and Comments:

In order to obtain exact birth date, this variable must be used in conjunction with DAYMAD and MONMAD. YERMAD appears on the PAR000 file as ANONAC for person number 02.

YERPAD Birth year of 01.

Input Variables:

425 Census, 100 series cards, columns 56-57.

Coding:

See YERMAD above for code.

Notes and Comments:

In order to obtain exact birth date, this variable must be used in conjunction with DAYPAD and MONPAD. YERPAD appears on the PAR000 file as ANONAC for person number 01.

YHMCON Value in quetzales for 6 major agricultural crops which were grown but not sold by the nuclear family in 1974.

Input Variables:

R10 survey, 300 series cards, columns 24-26 and 29-31.

Coding:

YHMCON is the value of

$$\sum_{i=1}^n (\text{HARVEST}_i - \text{SALES}_i) \text{PV}_i$$

where HARVEST_i = amount of crop i harvested
($i = 1, 2, \dots, 6$)

SALES_i = amount of crop i sold ($i = 1, 2, \dots, 6$)

PV_i = village price of crop i (see below
for prices used)

i = code number of major crop

1 = corn	4 = chile
2 = beans	5 = chile
3 = tomato	6 = barley

HARVEST was formed by summing amount harvested (columns 24-26) for each product after converting to a standard unit. SALES was formed by summing amount sold (columns 29-31) for each product after converting to standard units. See Table below for village crop prices used.

Table 5

VILLAGE PRICES OF SIX MAJOR CROPS BASED ON R10 FORM

Village	Crop					
	Corn	Beans	Tomato	Chile*	Chile**	Maicillo
03	Q 6.30	Q 16.65	Q 2.50	Q 6.00	Q 6.00	Q 5.05
06	Q 5.45	Q 17.00	Q 2.50	Q 3.00	Q 3.00	Q 4.50
08	Q 5.50	Q 17.00	Q 2.30	Q 6.10	Q 6.10	Q 5.00
14	Q 5.95	Q 16.55	Q 2.50	Q 7.00	Q 7.00	Q 5.60
31	Q 6.50	Q 15.75	Q 2.80	Q 6.20	Q 6.20	Q 5.00
32	Q 6.50	Q 15.75	Q 2.80	Q 6.20	Q 6.20	Q 5.00

* Chile Jalapeño: There are no chiles jalapeños produced in these communities.

** Chile Serrano

APPENDIX E

SAS CONTENTS OF THE DATA FILES

CONTENTS OF SAS DATA SET PAR000.SAS2

TRACKS USED=11 SUBEXTENTS=1 OBSERVATIONS=3198 CREATED BY JOB B9910UPD AT 11:05 MONDAY, NOVEMBER 17, 1980

BY SAS RELEASE 79.3A DSNAME=C.C5015.A3218.PAR000.SAS2 BLKSIZE=13030 LRECL=40 OBSERVATIONS PER TRACK=325 GENERATED BY DATA

LIST OF VARIABLES BY POSITION

#	VARIABLE	TYPE	LENGTH	POSITION	FORMAT	INFORMAT	LABEL
1	VILAGE	NUM	2	4			
2	NUCFAM	NUM	3	6			
3	PERSON	NUM	2	9			
4	SEXCEN	NUM	2	11			
5	ESTADO	NUM	2	13			
6	DAYNAC	NUM	2	15			
7	MESNAC	NUM	2	17			
8	ANONAC	NUM	2	19			
9	ALFAMA	NUM	2	21			
10	GRADOM	NUM	2	23			
11	OCPRNM	NUM	2	25			
12	OCACTM	NUM	2	27			
13	ALFAPA	NUM	2	29			
14	GRADOP	NUM	2	31			
15	OCPRNP	NUM	2	33			
16	OCACTP	NUM	2	35			
17	AGEMOS	NUM	3	37			

ALPHABETIC LIST OF VARIABLES

#	VARIABLE	TYPE	LENGTH	POSITION	FORMAT	INFORMAT	LABEL
17	AGEMOS	NUM	3	37			
9	ALFAMA	NUM	2	21			
13	ALFAPA	NUM	2	29			
8	ANONAC	NUM	2	19			
6	DAYNAC	NUM	2	15			
5	ESTADO	NUM	2	13			
10	GRADOM	NUM	2	23			
14	GRADOP	NUM	2	31			
7	MESNAC	NUM	2	17			
2	NUCFAM	NUM	3	6			
12	OCACTM	NUM	2	27			
16	OCACTP	NUM	2	35			
11	OCPRNM	NUM	2	25			
15	OCPRNP	NUM	2	33			
3	PERSON	NUM	2	9			
4	SEXCEN	NUM	2	11			
1	VILAGE	NUM	2	4			

----- SOURCE STATEMENTS -----
 |DATA OUT1.CLARK;
UPDATE INTER NEWCARDS; BY VILAGE NUCFAM PERSON;

CONTENTS OF SAS DATA SET CENNUC.SAS5

TRACKS USED=26 SUBEXTENTS=1 OBSERVATIONS=1829 CREATED BY JOB B9910NUC AT 8:54 TUESDAY, NOVEMBER 18, 1980
BY SAS RELEASE 79.3A DSNAME=C.C5015.A3218.CENNUC.SAS5 BLKSIZE=13030 LRECL=169 OBSERVATIONS PER TRACK=77
GENERATED BY PROC SORT

LIST OF VARIABLES BY POSITION

#	VARIABLE	TYPE	LENGTH	POSITION	FORMAT	INFORMAT	LABEL
1	VILAGE	NUM	2	4			
2	NUCFAM	NUM	3	6			
3	DADPRS	NUM	2	9			
4	MOMPRS	NUM	2	11			
5	PARPRS	NUM	2	13			
6	MB0000	NUM	2	15			
7	FB0000	NUM	2	17			
8	MB0106	NUM	2	19			
9	FB0106	NUM	2	21			
10	MB0710	NUM	2	23			
11	FB0710	NUM	2	25			
12	MB1114	NUM	2	27			
13	FB1114	NUM	2	29			
14	MB1545	NUM	2	31			
15	FB1545	NUM	2	33			
16	MB4659	NUM	2	35			
17	FB4659	NUM	2	37			
18	MB60UP	NUM	2	39			
19	FB60UP	NUM	2	41			
20	MT0000	NUM	2	43			
21	FT0000	NUM	2	45			
22	MT0106	NUM	2	47			
23	FT0106	NUM	2	49			
24	MT0710	NUM	2	51			
25	FT0710	NUM	2	53			
26	MT1114	NUM	2	55			
27	FT1114	NUM	2	57			
28	MT1545	NUM	2	59			
29	FT1545	NUM	2	61			
30	MT4659	NUM	2	63			
31	FT4659	NUM	2	65			
32	MT60UP	NUM	2	67			
33	FT60UP	NUM	2	69			
34	CNTPRF	NUM	2	71			
35	AGEMIS	NUM	2	73			
36	RESPDT	NUM	2	75			
37	SEXRES	NUM	2	77			
38	DAYCEN	NUM	2	79			
39	MONTHC	NUM	2	81			
40	YEARCE	NUM	2	83			
41	ADDRES	NUM	4	85			
42	ESTCIV	NUM	2	89			
43	ETHNIC	NUM	2	91			
44	RELSRA	NUM	2	93			

45	RELSNR	NUM	2	95
46	MEMBRZ	NUM	2	97
47	NHIJOS	NUM	2	99
48	NHJERT	NUM	2	101
49	NHORTI	NUM	2	103
50	NPERFM	NUM	2	105
51	NPERVI	NUM	2	107
52	PRGNOW	NUM	2	109
53	NURSNB	NUM	2	111
54	RELATN	NUM	2	113
55	HHFORM	NUM	2	115
56	EXTFAM	NUM	3	117
57	EXTNUM	NUM	2	120
58	NUMBPA	NUM	2	122
59	AGEM01	NUM	3	124
60	AGEM02	NUM	3	127
61	HIGRDP	NUM	2	130
62	GRAD01	NUM	2	132
63	GRAD02	NUM	2	134
64	ALFA01	NUM	2	136
65	ALFA02	NUM	2	138
66	OCCA01	NUM	2	140
67	OCCA02	NUM	2	142
68	OCCP01	NUM	2	144
69	OCCP02	NUM	2	146
70	CNTPIF	NUM	2	148
71	NUM006	NUM	2	150
72	NUM714	NUM	2	152
73	DENOMR	NUM	2	154
74	EACONF	NUM	3	156
75	AGEMSP	NUM	2	159
76	SEXMSF	NUM	2	161
77	DR006F	NUM	3	163
78	DR714F	NUM	3	166

ALPHABETIC LIST OF VARIABLES

#	VARIABLE	TYPE	LENGTH	POSITION	FORMAT	INFORMAT	LABEL
41	ADDRES	NUM	4	85			
59	AGEM01	NUM	3	124			
60	AGEM02	NUM	3	127			
35	AGEMIS	NUM	2	73			
75	AGEMSP	NUM	2	159			
64	ALFA01	NUM	2	136			
65	ALFA02	NUM	2	138			
70	CNTPIF	NUM	2	148			
34	CNTPRF	NUM	2	71			
3	DADPRS	NUM	2	9			
38	DAYCEN	NUM	2	79			
73	DENOMR	NUM	2	154			
77	DR006F	NUM	3	163			
78	DR714F	NUM	3	166			
74	EACONF	NUM	3	156			

42	ESTCIV	NUM	2	89
43	ETHNIC	NUM	2	91
56	EXTFAM	NUM	3	117
57	EXTNUM	NUM	2	120
7	PB0000	NUM	2	17
9	PB0106	NUM	2	21
11	PB0710	NUM	2	25
13	PB1114	NUM	2	29
15	PB1545	NUM	2	33
17	PB4659	NUM	2	37
19	PB60UP	NUM	2	41
21	PT0000	NUM	2	45
23	PT0106	NUM	2	49
25	PT0710	NUM	2	53
27	PT1114	NUM	2	57
29	PT1545	NUM	2	61
31	PT4659	NUM	2	65
33	PT60UP	NUM	2	69
62	GRAD01	NUM	2	132
63	GRAD02	NUM	2	134
55	HHPFORM	NUM	2	115
61	HIGRDP	NUM	2	130
6	MB0000	NUM	2	15
8	MB0106	NUM	2	19
10	MB0710	NUM	2	23
12	MB1114	NUM	2	27
14	MB1545	NUM	2	31
16	MB4659	NUM	2	35
18	MB60UP	NUM	2	39
4	MONPRS	NUM	2	11
39	MONTHC	NUM	2	81
20	MT0000	NUM	2	43
22	MT0106	NUM	2	47
24	MT0710	NUM	2	51
26	MT1114	NUM	2	55
28	MT1545	NUM	2	59
30	MT4659	NUM	2	63
32	MT60UP	NUM	2	67
46	NEMBRZ	NUM	2	97
47	NHIJOS	NUM	2	99
49	NMORTI	NUM	2	103
48	NMUERT	NUM	2	101
50	NPERFM	NUM	2	105
51	NPERVI	NUM	2	107
2	NUCFAM	NUM	3	6
71	NUM006	NUM	2	150
72	NUM714	NUM	2	152
58	NUMBFA	NUM	2	122
53	NURSNB	NUM	2	111
66	OCCA01	NUM	2	140
67	OCCA02	NUM	2	142
68	OCCP01	NUM	2	144
69	OCCP02	NUM	2	146
5	PARPRS	NUM	2	13

52	PRGNO	NUM	2	109
54	RELATN	NUM	2	113
45	RELSNR	NUM	2	95
44	RELSRA	NUM	2	93
36	BESPDT	NUM	2	75
76	SEXHSF	NUM	2	161
37	SEXRES	NUM	2	77
1	VILAGE	NUM	2	4
40	YEARCE	NUM	2	83

CONTENTS OF SAS DATA SET CENEXT.SAS2

TRACKS USED=8 SUBEXTENTS=1 OBSERVATIONS=1553 CREATED BY JOB B9910CEN AT 15:07 MONDAY, NOVEMBER 17, 1980
 BY SAS RELEASE 79.3A DSNAME=C.C5015.A3218.CENEXT.SAS2 BLKSIZE=13030 LRECL=44 OBSERVATIONS PER TRACK=296
 GENERATED BY PROC SORT

LIST OF VARIABLES BY POSITION

#	VARIABLE	TYPE	LENGTH	POSITION	FORMAT	INFORMAT	LABEL
1	VILAGE	NUM	2	4			
2	EXTFAM	NUM	3	6			
3	CNTFIX	NUM	2	9			
4	HHMODE	NUM	2	11			
5	HIGRDX	NUM	2	13			
6	CENPIV	NUM	2	15			
7	AGEMIX	NUM	2	17			
8	CENPIX	NUM	2	19			
9	CNT1PX	NUM	2	21			
10	CNT2PX	NUM	2	23			
11	AGEMSX	NUM	2	25			
12	DENOMX	NUM	2	27			
13	NM006X	NUM	2	29			
14	NM714X	NUM	2	31			
15	SEXMSX	NUM	2	33			
16	EACONX	NUM	3	35			
17	DR006X	NUM	3	38			
18	DR714X	NUM	3	41			

ALPHABETIC LIST OF VARIABLES

#	VARIABLE	TYPE	LENGTH	POSITION	FORMAT	INFORMAT	LABEL
7	AGEMIX	NUM	2	17			
11	AGEMSX	NUM	2	25			
6	CENPIV	NUM	2	15			
8	CENPIX	NUM	2	19			
3	CNTFIX	NUM	2	9			
9	CNT1PX	NUM	2	21			
10	CNT2PX	NUM	2	23			
12	DENOMX	NUM	2	27			
17	DR006X	NUM	3	38			
18	DR714X	NUM	3	41			
16	EACONX	NUM	3	35			
2	EXTFAM	NUM	3	6			
4	HHMODE	NUM	2	11			
5	HIGRDX	NUM	2	13			
13	NM006X	NUM	2	29			
14	NM714X	NUM	2	31			
15	SEXMSX	NUM	2	33			
1	VILAGE	NUM	2	4			

CONTENTS OF SAS DATA SET INCOME10.SAS3

TAPE FORMAT DATA SET CREATED BY JOB B9910DTP AT 15:53 MONDAY, NOVEMBER 17, 1980 BY SAS RELEASE 79.3A DSNAME=INCOME10.SAS3

BLKSIZE=20480 LRECL=243 GENERATED BY PROC COPY

ALPHABETIC LIST OF VARIABLES

#	VARIABLE	TYPE	LENGTH	POSITION	FORMAT	INFORMAT	LABEL
19	CINRST	NUM	3	57			
71	CIVMAD	NUM	3	210			
60	CIVPAD	NUM	3	177			
41	CIVSTA	NUM	3	120			
13	CLABOR	NUM	3	39			
14	CLANDP	NJM	3	42			
77	CLSMAD	NUM	3	228			
66	CLSPAD	NUM	3	195			
38	CONDUR	NUM	3	113			
40	CONSUP	NUM	2	118			
52	COOKNG	NUM	3	153			
10	CVARIN	NUM	3	30			
72	DAYMAD	NUM	3	213			
61	DAYPAD	NUM	3	180			
57	DXHOME	NUM	3	168			
58	DXUTLY	NUM	3	171			
39	EXSTRC	NUM	2	116			
76	GRAMAD	NUM	3	225			
65	GRAPAD	NUM	3	192			
51	KITCHN	NUM	3	150			
53	LIGHTS	NUM	3	156			
75	LITHAD	NUM	3	222			
64	LITPAD	NUM	3	189			
48	MFLOOR	NJM	3	141			
80	HIGHAD	NUM	3	237			
69	HIGPAD	NUM	3	204			
73	HONHAD	NUM	3	216			
62	HONPAD	NUM	3	183			
49	HTROOF	NUM	3	144			
50	MTWALL	NUM	3	147			
59	NODURB	NUM	3	174			
47	NOROOM	NUM	3	138			
2	NUCFAM	NUM	3	6			
79	OCACTM	NUM	3	234			
68	OCACTP	NUM	3	201			
78	OCPRNM	NJM	3	231			
67	OCPRNP	NUM	3	198			
46	OWNHME	NUM	3	135			
44	OWNPRO	NUM	3	129			
42	PERFAM	NUM	3	123			
43	PERHSE	NUM	3	126			
36	PRODUR	NUM	3	107			
81	PROMAD	NUM	3	240			
70	PROPAD	NUM	3	207			
9	SANIML	NUM	3	27			

S T A T I S T I C A L A N A L Y S I S S Y S T E M

10:51 THURSDAY, NOVEMBER 20, 1980 ²

54	SANITN	NUM	3	159
55	SEWDSP	NJM	3	162
3	SPADAD	NUM	3	9
4	SPANOM	NUM	3	12
5	SPAOT1	NUM	3	15
6	SPAOT2	NUM	3	18
30	SSORCH	NUM	2	90
32	SSORCW	NJM	2	94
56	SSWATR	NUM	3	165
37	TRADUR	NUM	3	110
22	TRNOME	NUM	3	66
23	TRNREG	NUM	3	69
21	TRWTOT	NUM	3	63
45	TYPHME	NUM	3	132
1	VILAGE	NUM	2	4
34	VLANDN	NUM	4	99
35	VLANDX	NUM	4	103
33	WRKWIF	NUM	3	96
20	YAGPRO	NUM	3	60
8	YDERPR	NUM	3	24
74	YERHAD	NUM	3	219
63	YERPAD	NUM	3	186
12	YHMCON	NUM	3	36
18	YINRST	NJM	3	54
15	YLANDR	NUM	3	45
24	YNALLT	NUM	3	72
25	YNNOTR	NJM	3	75
27	YNPALT	NUM	3	81
28	YNPNOT	NUM	3	84
29	YNPRGT	NJM	3	87
26	YNREGT	NUM	3	78
7	YPAALL	NUM	3	21
16	YRENTD	NUM	3	48
17	YSALED	NUM	3	51
11	YSCROP	NUM	3	33
31	YSORCH	NUM	2	92

CCNTENTS OF SAS DATA SET INCEXT.SAS1

TRACKS USED=6 SUBEXTENTS=1 OBSERVATIONS=1554 CREATED BY JOB B9910MER AT 11:08 THURSDAY, NOVEMBER 20, 1980

BY SAS RELEASE 79.3A DSNAME=C.C5015.A3218.INCEXT.SAS1 BLKSIZE=19069 LRECL=83 OBSERVATIONS PER TRACK=229 GENERATED BY DATA

ALPHABETIC LIST OF VARIABLES

#	VARIABLE	TYPE	LENGTH	POSITION	FORMAT	INFORMAT	LABEL
22	ACNDRX	NUM	3	62			AVE CONSUMER DURABLE INDEX PER NUC FAMLY
24	ALANDX	NUM	3	68			AVE OF RESPONSES FOR EXTFAM LAND VALUE
23	APKDRX	NUM	3	65			AVE PRODUCER DURABLE INDEX PER NUC FAMLY
6	CLABRX	NUM	3	15			TOTAL LABOR COSTS FOR EXTENDED FAMILY
4	CNTFX	NUM	2	11			TALLY OF NUCLEAR FAMILIES IN EXTFAM
17	CNT2XX	NUM	2	48			TALLY OF PEOPLE IN EXTENDED FAMILY
18	EACNXX	NUM	3	50			EXTFAM EQUIVALENT ADULT CONSUMERS (EAC)
2	EXTFAM	NUM	3	6			EXTENDED FAMILY FOR 74 FROM CENNUC FILE
3	MSEACX	NUM	2	9			EACNXX CODE: 0 MEANS NO AGE DATA MISSING
28	MXLNDX	NUM	3	80			MAX OF RESPONSES FOR EXTFAM LAND VALUE
7	SCNDRX	NUM	3	18			TOTAL CONSUMER DURABLE INDEX FOR EXTFAM
9	SLANDX	NUM	3	24			TOTAL OF RESPONSES FOR EXTFAM LAND VALUE
8	SPRDRX	NUM	3	21			TOTAL PRODUCER DURABLE INDEX FOR EXTFAM
1	VILAGE	NUM	2	4			LARGE ARE 3 AND 6, ATCLE ARE 6 AND 14
10	YAGPRX	NUM	3	27			NET EXTFAM INCOME FROM AGRI PRODUCTION
11	YDRPRX	NUM	3	30			NET EXTFAM INCCME FROM DERIVED PRODUCTS
5	YFORMX	NUM	2	13			TYPE OF HCUSEHOLD(NUCLEAR,EXTENDED,ETC)
12	YHMCNX	NUM	3	33			EXT FAMILY INCCME FROM HOME CONSUMPTION
13	YNALTX	NUM	3	36			NET EXTENDED FAMILY INCOME(ALL TRANSFERS)
19	YNEXAT	NUM	3	53			NET EXTFAM INCOME PER EAC(ALL TRANSFERS)
20	YNEXNT	NUM	3	56			NET EXTFAM INCOME PER EAC(NO TRANSFERS)
21	YNEXRT	NUM	3	59			NET EXTFAM INCOME PER EAC(REG TRANSFERS)
14	YNOTX	NUM	3	39			NET EXTENDED FAMILY INCOME(NO TRANSFERS)
25	YNPXAT	NUM	3	71			NET EXTFAM INCCME PER CAPITA(ALL TRNSFR)
26	YNPXNT	NUM	3	74			NET EXTFAM INCCME PER CAPITA(NC TRNSFRS)
27	YNPXRT	NUM	3	77			NET EXTFAM INCOME PER CAPITA(REG TRNSFR)
15	YNRGTX	NUM	3	42			NET EXTENDED FAMILY INCOME(REG TRANSFERS)
16	YPAALX	NUM	3	45			INCOME FROM SALARY FOR EXTENDED FAMILY

BIBLIOGRAPHY

- Church, Charles and Helen Church, Food Value of Portions Commonly Used, L. P. Lippincott Company, New York, 1975.
- Corona, Henry L., Codebook and User's Manual: INCAP-Rand Guatemala Survey, The Rand Corporation, N-1566, August 1978.
- Helwig, Jane T., and Kathryn A. Council, SAS User's Guide 1979 Edition, SAS Institute Inc., North Carolina, 1979.
- Instituto de Nutrición de Centro América y Panamá, Recomendaciones Dietéticas Diarias para Centro América y Panamá, INCAP Publication E-709, Guatemala, 1973.
- Kasala, Cathy, Documentation of Programs and Data Sets from the Income, Wealth, and Agricultural Production Questionnaire (R10) of the Rand/INCAP Guatemala Project. The Rand Corporation, WD-789, May 1979.
- Sukhatme, P. V., "Assessment of Food and Nutrition," in Donald S. McLaren (ed.), Nutrition in the Community, John Wiley and Sons, New York, 1976.

